AD-A100 138   PITTSBURGH UNIV PA LEARNING RESEARCH AND DEVELOPMEN--ETC F/G 5/10
EXPERTISE IN PROBLEM SOLVING.(U)
MAY 81 M T CHI, R GLASER, E REES   N00014-78-C-0375
UNCLASSIFIED TR-5   NL

1 OF 2
AD A
100138

LEVEL II

(12)

University of Pittsburgh

LEARNING RESEARCH AND DEVELOPMENT CENTER

EXPERTISE IN PROBLEM SOLVING

Michelene T.H. Chi, Robert Glaser, and Ernest Rees
Learning Research and Development Center
University of Pittsburgh

18 May 1981

Technical Report No. 5

This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-78-C-0375, Contract Authority Identification Number, NR 157-421.

This report is issued by the Learning Research and Development Center, supported in part by funds from the National Institute of Education (NIE), United States Department of Education.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE | READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Technical Report No. 5	2. GOVT ACCESSION NO. AD-A100 138	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) EXPERTISE IN PROBLEM SOLVING		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Michelene T. H. Chi, Robert Glaser and Ernest Rees		8. CONTRACT OR GRANT NUMBER(s) N00014-78-C-0375
9. PERFORMING ORGANIZATION NAME AND ADDRESS Learning Research and Development Center University of Pittsburgh Pittsburgh, PA 15260		10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS NR 157-421/3-3-78 (458)
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Programs Office of Naval Research (Code 458) Arlington, VA 22217		12. REPORT DATE 18 May 1981
		13. NUMBER OF PAGES 136
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES This research was also supported by the Learning Research and Development Center, supported in part by funds from the National Institute of Education (NIE), United States Department of Education. To be published in R. Sternberg (Ed.), Advances in the psychology of human intelligence. Hillsdale, N.J.: Erlbaum, 1981.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Production systems
Means-ends
Knowledge
Problem solving
Representation
Chess
Physics

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It has become increasingly clear in recent years that the quality of domain-specific knowledge is the main determinant of expertise in that domain. This paper begins with an examination of the shift from consideration of general, domain-independent skills and procedures, in both cognitive psychology and artificial intelligence, to the study of the knowledge base. Next, the empirical findings and theoretical models of other researchers in physics problem solving are detailed and summarized.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. Then our own work is presented, consisting of eight empirical studies. These studies show, in general, the importance of differences in the knowledge bases of experts and novices to their problem solving success. More specifically, they show (1) that it is difficult to use protocols of problem solving episodes to illuminate the differences in the knowledge bases of experts and novices, (2) that experts and novices perceive the problem themselves differently, i.e., novices respond to the surface features of a problem while experts respond to its deep structure, (3) that less successful novices, at least, have deficiencies in their declarative knowledge of physics, (4) that novices tend to lack knowledge of when to use certain physics knowledge, and (5) that deficiencies in knowledge appear to prevent novices at times from making key inferences necessary for solving problems. Finally, these results and their implications for theories of intelligence, are discussed.

Abstract

It has become increasingly clear in recent years that the quality of domain-specific knowledge is the main determinant of expertise in that domain. This paper begins with an examination of the shift from consideration of general, domain-independent skills and procedures, in both cognitive psychology and artificial intelligence, to the study of the knowledge base. Next, the empirical findings and theoretical models of other researchers in physics problem solving are detailed and summarized. Then our own work is presented, consisting of eight empirical studies. These studies show, in general, the importance of differences in the knowledge bases of experts and novices to their problem solving success. More specifically, they show (1) that it is difficult to use protocols of problem solving episodes to illuminate the differences in the knowledge bases of experts and novices, (2) that experts and novices perceive the problems themselves differently, i.e., novices respond to the surface features of a problem while experts respond to its deep structure, (3) that less successful novices, at least, have deficiencies in their declarative knowledge of physics, (4) that novices tend to lack knowledge of when to use certain physics knowledge, and (5) that deficiencies in knowledge appear to prevent novices at times from making key inferences necessary for solving problems. Finally, these results and their implications for theories of intelligence are discussed.

Expertise in Problem Solving

INTRODUCTION

At first glance, it may seem anomalous for a chapter on expert performance to appear in a volume on intelligence. But an accumulation of scientific events indicates that the analysis of expertise in semantically rich knowledge domains is quite relevant to understanding the nature of intelligence. These events have occurred in a number of disciplines, particularly cognitive psychology and artificial intelligence. The first part of this paper briefly outlines work in these fields. The common theme is the increasing emphasis on the structure of knowledge as a significant influence on intelligence and high level cognitive performance. The latter part of this paper describes, as an illustration of this, investigations of high and low competence in a knowledge-rich domain, namely, problem solving in physics.

Intelligence has been studied by contrasting individual differences, age differences, differences between the retarded and the gifted, and between fast and slow learners. These dimesions of difference are well represented by the past research of the contributors to this volume, including ourselves. What have we learned by investigating intelligent performance along these dimensions? If we consider speed of processing, memory span, and the

use of complex strategies as three straightforward measures of cognitive performance, the following picture emerges. More intelligent individuals have faster processing speed, longer memory span, and use more sophisticated strategies than less intelligent persons (Belmont & Butterfield, 1971; Hunt, Lunneborg, & Lewis, 1975; Jenson, in press). This is also true of older versus younger children (Chi, 1976), and fast as compared with slow learners. For example, good readers can encode words faster and have a longer memory span for words than poor readers (Perfetti & Hogaboam, 1975). Thus, over these dimensions of comparison, measured intelligence correlates positively with faster processing, more complex encoding and recall, and the use of sophisticated strategies.

Although this pattern of results occurs reliably, we still do not understand what the underlying mechanisms are, and whether similar mechanisms are operative in various disciplines and areas of knowledge. This is one reason the analysis of expertise has emerged as an interesting area of investigation. The study of expertise forces us to focus on a new dimension of difference between more and less intelligent individuals--the dimension of knowledge--since expertise is, by definition, the possession of a large body of knowledge and procedural skill. The central thesis of this paper is that a major component of intelligence is the possession of a large body of accessible and usable knowledge. In the following section, we briefly outline the literature in two related disciplines that have gradually come to the same conclusion.

THE FOCUS ON KNOWLEDGE

Cognitive Psychology

Memory Skills

In cognitive psychology, the effects of knowledge on complex skilled performance were first explored in the seminal work of de Groot (1966) and Chase and Simon (1973a, 1973b) in their studies of chess skill. In an attempt to discover what constitutes skill in chess, de Groot (1966) found that differences in skill were not reflected in the number of moves the players considered during their search for a good move, nor in the depth of their search. Both the master and the novice did not search any further ahead than five moves. Both experts and novices used the same search strategies, that is, depth first with progressing deepening. In order to capture the essence of skill differences in chess, de Groot resorted to a different type of task--memory for chess positions. He found that when masters were shown a chess position for a very brief duration (five seconds), they were able to remember the position far better than the novice players. This difference could not be attributed to superior visual short-term memory on the part of the masters because, when random board positions were used, recall was equally poor for masters and novices (Chase & Simon, 1973a).

In order to understand the chess masters' recall superiority, Chase and Simon attempted to uncover the structures of chess knowledge that the masters possessed. Using <u>chunks</u> as a defining unit of knowledge structure, Chase and Simon set out to experimentally identify the structure and size of chunks in the knowledge base of

masters and novices. Two procedures were used by Chase and Simon. One was to record the placement of chess pieces on the chess board during the recall of positions, and use two-second pauses during recall to segment the chunks. A second procedure was asking the chess player to copy a position and using head turns from board to board to partition the chunks. The theoretical rationale underlying both the pause and the head-turn procedure was the notion that chunks are closely knit units of knowledge structure; hence, retrieval of one item of information within a chunk would lead to retrieval of another in quick succession.

Both master and novice did retrieve pieces in chunks--bursts followed by pauses, and they reproduced chess positions pattern by pattern, with a glance (or head turn) for each pattern. These patterns were familiar and highly stereotypic patterns that chess players see daily, such as a castled-king position, or a pawn chain, or they were highly circumscribed clusters of pieces, often of the same color, and located in very close proximity. The difference between the novice and the expert chess player was the size of the chunks. The master's patterns were larger, containing three to six pieces, whereas novice's patterns contained single pieces. If one counted by chunks rather than pieces, the novice and the master were recalling the same number of chunks from the board position.

There are limitations with the procedure of identifying chunks by a two-second pause and/or a head turn. One limitation is that it does not provide a description of the complex structure of the chunk, for example, the overlapping nature of chunks (Reitman, 1976). A more

serious limitation is that it does not allow for the identification of higher-order chunks. The pause procedure permits only the identification of "local" chunks, that is, chunks that are spatially close and defined by such relations as next to, color identity, piece identity, etc. (Chase & Chi, in press).

The existence of higher-order chunks is evidenced in the master's recall for sequences of moves (Chase & Simon, 1973b). That is, after viewing all the moves of a game, a master's recall of move sequences shows clustering of move sequences represented by pauses that is similar to the clustering of pieces in the board-recall task. This says that a given board position generates a sequence of stereotypic moves. Data from eye movement studies clearly show that chess players fixate predominantly on the pieces interrelated by attack and defense strategy (Simon & Barenfeld, 1969), and that these pieces are typically not proximally related, as are the local chunk pieces.

The study of expert-novice differences in the use of complex knowledge in other domains has also revealed higher-order chunk structures. In electronics, Egan and Schwartz (1979) found that skilled technicians reconstructing symbolic drawings of circuit diagrams do so according to the functional nature of the elements in the circuit such as amplifiers, rectifiers, and filters. Novice technicians, however, produce chunks based more upon the spatial proximity of the elements. In architecture, Akin (1980) found that during recall of building plans by architects, several levels of patterns were produced. First, local patterns consisting of wall segments and doors are recalled, then rooms and other areas, then

clusters of rooms or areas. The hierarchical nature of chunks also has been illustrated in the recall of baseball events. High-knowledge individuals can recall entire sequences of baseball events much better than low-knowledge individuals (Chiesi, Spilich, & Voss, 1979).

Like the chess results, the expert in several diverse domains is able to remember "sequences of moves" much more rapidly than novices. Also, we see a similarity between chess patterns, circuit diagrams, and architectural patterns in that functional properties are more important at higher levels, whereas structural properties (such as proximity and identity in color and form) are more important at lower levels. And with increasing skill, more higher-order chunks are developed.

In sum, one aspect of cognitive psychology research has clearly identified the superior memory capacity of skilled individuals, as exhibited in the large pattern of chunks, whether they are adult chess players, child chess players (Chi, 1978), Go players (Reitman, 1976), Gomoku players (Eisenstadt & Kareev, 1975), bridge players (Charness, 1979), musicians (Sloboda, 1976), baseball fans (Chiesi, Spilich, & Voss, 1979), programmers (McKeithen, 1979; Jeffries, Turner, Polson, & Atwood, 1981), or electronic technicians (Egan & Schwartz, 1979). While a number of the above studies have uncovered the hierarchical nature of the patterns (Akin, 1980; Chiesi, Spilich, & Voss, 1979; Egan & Schwartz, 1979), no work to date has explicitly related the knowledge and chunk structures of these skilled individuals to the complex skill that they are able to perform.

Problem-Solving Skills

A currently prominent area of research in cognitive psychology is problem solving. Problem-solving research was revolutionized in the sixties when researchers turned from studying the conditions under which solutions are reached to the processes of problem solving. Following the contribution of Newell and Simon's (1972) theory, problem-solving research proceeded to model search behavior, and to verify that humans indeed solve problems according to means-ends analyses. Numerous puzzle-like problems were investigated, all of which indicated that human subjects do solve problems according to means-ends analyses to some degree (Greeno, 1978).

In puzzle problems, sometimes known as MOVE problems, the knowledge involved in solving the problems is minimal. All the knowledge one needs to solve the problems is given: the initial state, the number and function of operators, and the final goal state. Solution requires that a set of operators be applied to transform one state of knowledge to another, so that eventually the goal state can be reached. A variety of puzzle problems have been investigated: the water jug problem (Atwood & Polson, 1976; Atwood, Masson, & Polson, 1980; Polson & Jeffries, this volume), hobbits and orcs (Greeno, 1974; Thomas, 1974), missionaries and cannibals (Reed & Simon, 1976), and Tower of Hanoi (Egan & Greeno, 1974; Simon, 1975).

The research on puzzle problems, however, offered limited insights into learning. Because, learning in real-world subject matters requires the acquisition of large bodies of domain-specific knowledge, cognitive scientists turned their attention from

knowledge-free problems, like puzzles, to knowledge-filled domains like geometry (Greeno, 1978), physics (Simon & Simon, 1978), thermodynamics (Bhaskar & Simon, 1977), programming (Polson, 1981), understanding electronic circuits (Brown, Collins, & Harris, 1978), and recently, political science (Voss & Tyler, 1981).

Solving real-world problems presents new obstacles that were not encountered previously in puzzle-like problems. Basically, the exact operators to be used are usually not given, the goal state is sometimes not well defined, and more importantly, search in a large knowledge space becomes a serious problem. (The research on artificial intelligence programs in chess, to be mentioned in the next section, gives the flavor of this difficulty.) Solving real-world problems with large knowledge bases also provides a glimpse of the power of the human cognitive system to use a large knowledge system in an efficient and automatic manner--in ways that minimize heuristic search. In general, current studies of high levels of competence by cognitive psychologists appear to support the recommendation that a significant focus for understanding expertise is investigation of the characteristics and influence of organized, hierarchical knowledge structures that are acquired over years of learning and experience.

Artificial Intelligence

The goal of artifical intelligence (AI) research is to make a machine act intelligently. In this area, the problem of understanding intelligence has become increasingly focused on the large structure of domain-specific knowledge that is characteristic of experts. This is

in contrast to the early years of the field, when the creation of intelligent programs was identified with finding "pure" problem-solving techniques to guide a search, for any problem, through the problem space to a solution, as in the General Problem Solver (Newell, Shaw, & Simon, 1960). The techniques elucidated, such as means-end analysis, are clearly part of the picture, but it was apparent early on that in realistically complex domains, such techniques must engage a highly organized structure of specific knowledge. This shift in AI is characterized by Minsky and Papert (1974) as a change from a power-based strategy for achieving intelligence to a knowledge-based emphasis. They write as follows:

> The Power strategy seeks a generalized increase in computational power. It may look toward new kinds of computers ("parallel" or "fuzzy" or "associative" or whatever) or it may look toward extensions of deductive generality, or information retrieval, or search algorithms....In each case the improvement sought is intended to be "uniform"--independent of the particular data base.
>
> The Knowledge strategy sees progress as coming from better ways to express, recognize, and use diverse and particular forms of knowledge. This theory sees the problem as epistemological rather than as a matter of computational power or mathematical generality. It supposes, for example, that when a scientist solves a new problem, he engages a highly organized structure of especially appropriate facts,

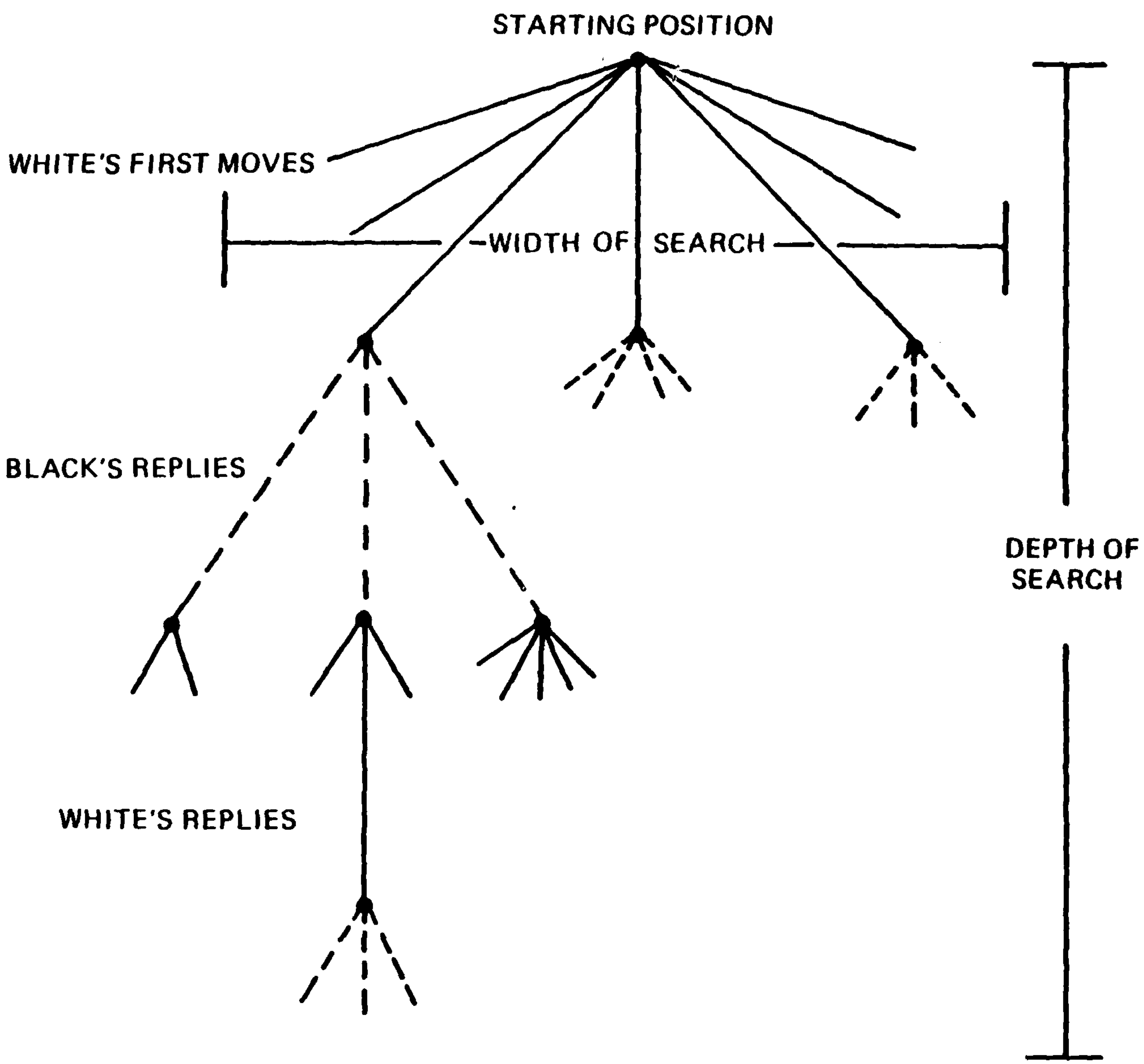


Figure 1. A chess game tree.

> models, analogies, planning mechanisms, self-discipline procedures, etc. To be sure, he also engages "general" problem-solving schemata but it is by no means obvious that very smart people are that way directly because of the superior power of their general methods--as compared with average people. Indirectly, perhaps, but that is another matter: A very intelligent person might be that way because of specific local features of his knowledge-organizing knowledge rather than because of global qualities of his "thinking" which, except for the effects of his self-applied knowledge, might be little different from a child's. (p. 59).

We can now elaborate on this transition in AI research from building programs that emphasized heuristic search to knowledge-based programs, using chess programs as examples. The chess problem space can be pictured as a game tree. Figure 1 shows a very simple example of such a tree. Each node represents a possible position (of all the pieces) during a game and each link leading from a node represents a possible move. At first glance, the problem might seem fairly simple: Start at the top of the tree and find a set of paths that force the opponent into checkmate. However, as Shannon (1950) pointed out, at any given point a player has on the order of 30 legal moves available, so the number of nodes at successive levels of the tree increases dramatically. In an entire game, each player makes an average of 40 moves (giving the tree 80 levels) and the number of possible paths to the bottom of the tree total about 10^{120}. Even the fastest computer

could not search such a tree exhaustively, so intelligent choices must be made to severely limit the exploration. There are two basic limitations that can be applied: limiting the number of moves considered from each node (width of search) and limiting the number of successive moves that will be considered on each path (depth of search). Both of these methods require some chess knowledge to be used if they are to be applied successfully. In the case of depth of search, since positions reached are not final (won or lost), they must be evaluated to determine if they are advantageous or not. In addition, simply cutting off the search at a specified depth can cause problems (for example, the cut off may be in the middle of an exchange of pieces), so some analysis is required to determine if the search should be deepened.

Full-Width Search

Two general search-based approaches have been followed in attempts to create chess playing programs: full-width (brute force) search and selective search. Both limit the depth of search. In a full-width program, as the name implies, the width of search is not limited at all. To date, a modification of this approach has been the most successful. It uses a mathematical algorithm which eliminates from consideration moves by the opponent which are worse than the best move already found (based on the evaluation of the positions to which they lead) since it must be assumed that he will make his best possible move. The current (1980) world computer chess champion, BELLE by Thompson and Condon at Bell Labs, and the former champion,

CHESS 4.6 by Slate and Atkin at Northwestern, are both of this type. These programs, and others like them, make use of a computer's speed and memory to do vast amounts of searching, and have a bare minimum of chess knowledge. Although these programs can now beat practically all human players, they cannot beat the top ranked experts (grandmasters). Estimates of 10 more years of work to reach this level are not uncommon. The main reason for such slow progress is probably the explosive branching of the game tree. Each level contains about 30 times as many nodes as the level above so a large increase in computing power is needed for a very small increase in depth of search.

Selective Search

Clearly, grandmasters do not play better chess because they outsearch a computer. The limited size of short-term memory and the amount of time required to fixate items in long-term memory limit humans to very tiny tree searches. In fact, de Groot (1965) and Newell and Simon (1972) have shown through protocol analysis that expert players tend to choose good moves without any search at all and then conduct a limited search to test their choices. This approach is an example of the second programming method--selective search. The Greenblatt (1967) program, the first to make a respectable showing in human tournament competition, provides an example of how this approach has been implemented. His program selects moves for consideration on the basis of "plausbility." It first generates all of the legal moves available from the present position. A plausibility score is then

calculated for each move on the basis of a subset of 50 heuristics (not all are applicable to a given situation). These heuristics are simply "rules of thumb" for selecting a good move, taken from chess lore, which have been roughly quantified to allow a numerical score to be calculated. The moves are then ranked in order of decreasing plausibility and only the first few are considered. In addition, all of the continuations used to evaluate a move are generated in the same way. Since only a handful of the possible moves is considered at each node, the game tree is significantly reduced in size. The size of the search must be reduced still further, however, so the mathematical algorithm mentioned before is used to "prune" more branches from the tree and the depth of search is also limited.

Although expert players do choose a few plausible moves for consideration, they do not do it through computation and evaluation as does the Greenblatt program. Rather, they respond intuitively to patterns on the board. As mentioned earlier, de Groot (1965) has shown that grandmasters can reproduce complicated positions almost exactly after seeing them for only five seconds. Apparently, the years of practice necessary to become a chess expert result in a very large knowledge base of patterns of pieces and probably patterns of moves as well. When an expert looks at the board and "sees" good moves, he is engaging in pattern recognition. Thus, an obvious direction for chess program design is to build production systems that can recognize and respond as human players do (Simon, 1976).

Knowledge-Based Chess

There is more to human play than just recognizing a possible next move, however. The moves of a good player advance toward some goal; they fit into a plan that looks at least a few moves ahead. An early attempt to give chess programs simple goals is the Newell, Shaw, and Simon program (1958). It has a series of independent goal modules. Each module can recognize appropriate situations on the board and generate moves with specific purposes, such as king safety, center control, etc. The purpose of these goals, however, is only to select a few reasonable candidates for the next move in order to limit the search tree; there is no overall plan.

A program called PARADISE, written by Wilkins (1980), contains the factors we have discussed that seem to give expert chess players an edge over even the best search programs. It uses an extensive knowledge of chess board patterns, embodied in production rules, to establish goals, which are then elaborated into more concrete plans. Search is used only to check the validity of the plans.

PARADISE does not play an entire game; it plays "tactically sharp" positions from the middle game. Tactically sharp simply means that success can be achieved by winning material from the opponent--a common situation in chess. The knowledge base consists of some 200 production rules; each has a general pattern of relationships among pieces as its condition. Most of these rules are organized around general higher level concepts necessary for effective play, such as looking for a THREAT to the opponent's pieces, looking for a way to make a square SAFE to move a piece to it, trying to DECOY an opponent

piece out of the way, etc. The effect of applying the production rules to a given position is to suggest a plan or plans with the overall goal of winning material. A given plan may include calls back to the knowledge base to produce plans to accomplish subgoals of the original plan (if such a subplan cannot be found, then the overall plan is scrapped). Plans are thus hierarchically expanded until they are ready for use. Each plan contains an initial move plus a series of alternative future moves depending on the types of replies by the opponent. Each plan also contains information about why the knowledge base produced it in the first place. The plan and its associated information are then used to guide a very small tree search to determine if the plan is feasible.

Productions in the knowledge base are used to generate the defensive moves used in the search. Calls for additional planning and analysis to expand the original plan can also be generated by the search. The depth of search is not artificially limited in this program; instead, analyses are conducted (using the knowledge base) at the ends of lines suggested by the plans to determine if termination of the search is proper. Since the plans limit the number of alternatives considered at each node to only a few, the search can go much deeper than in other programs. Since all of the analysis, planning, and searching is guided by the knowledge base, altering or improving the play of PARADISE consists of simply modifying or adding individual production rules. Such a system seems to have great potential for playing expert chess, if the requisite knowledge can be

determined and coded into the knowledge base, or if a self-learning system can be designed to modify its own base.

In sum, the example of chess programs illustrates the general tendency in AI toward knowledge-based programming. Even though computers have great advantages over humans in speed and memory, knowledge provides an edge which, it seems, pure power can only overcome at great cost, if at all.

PHYSICS PROBLEM SOLVING AND EXPERTISE

In this section, we review what is known about how physics problems are solved; and in particular, how expert physicists solve them as compared to novices. The first section reviews the available empirical evidence, and the second section reviews the resulting theoretical models simulating the way experts and novices solve physics problems.

Empirical Findings

In the relatively small amount of work done in this area, there are basically three types of empirical investigation. One is examination of the knowledge structures of physics concepts. Shavelson (1974, also Shavelson & Stanton, 1975) for instance, has investigated methods for determining this "cognitive structure." He delineates three methods that may be used singly or in conjunction: word association, card sorting, and graph building. Of the three, word association is the most venerable and widely used. Using this method, Shavelson (1974) has shown that students' physics concepts

become more interrelated and that their cognitive structures become more like the course "content structure" (as determined by a structural analysis of the instructional materials) at the end of the course than at the beginning. Thro (1978) has found similar results using the instructors' cognitive structure as the content structure.

A second type of empirical research is investigation of subjects' prior conception of the physical world, with a view toward how that preconception might affect one's learning of physics. For example, McCloskey, Caramazza, & Green (1980), have shown that a sizable number of students who have had no physics courses, as well as some who have had one or more college courses, believe that an object once set in curvilinear motion (through a spiral tube, for instance) will maintain that motion in the absence of any further external forces. Also, Champagne and Klopfer (1980) have constructed the Demonstration, Observation, and Explanation of Motion Test (D.O.E) to test students' ideas of motion due to gravity. They have found, similarly, that a sizable number of students entering a college mechanics course have erroneous ideas about motion (and that students who had taken high school physics did no better than those who had not). They also found, however, that results on the D.O.E. alone were of little predictive value in determining success in the mechanics courses.

The third type of empirical evidence relates specifically to problem solving and is usually gathered in the context of solution protocols. Careful analyses of protocols have indicated significant differences between the expert and novice. The only obvious similarities between them are in the macroprocesses they use in

solving physics problems. According to Simon and Simon (1978), both expert and novice proceed to solution by evoking the appropriate physics equations, and then solving them. Their expert often did this in one step, however, simply stating results without explicitly mentioning the formula he was using, while the novice typically stated the formula, put it into the appropriate form and substituted the values of the variables in discrete steps. McDermott and Larkin (1978) include another two "stages" prior to the evoking and instantiating of equations, postulating that solution proceeds in at least four episodes: the first "stage" is simply the written problem statement. The second "stage" involves drawing a sketch of the situation, while the third is a "qualitative analysis" of the problem which results in a representation containing abstract physics entities. Generating the equations is the fourth stage. According to Larkin (1980) experts seem to perform all four processes, whereas the novice may skip the "qualitative analysis" stage. Beyond this gross similarity lies much more subtle and salient differences between the expert and novice protocols. These are elaborated below.

<u>Quantitative Differences</u>

There are three major differences between the novice and the expert physicist that are easily quantifiable. The most obvious one is time to solution. The speed with which a problem can be solved depends a great deal on the skill of the individual. Simon and Simon (1978) noted a 4 to 1 difference between their expert and novice. Larkin (1981) also reported a similar difference between her experts

and novices. This difference is not unlike the speed difference found in chess-playing ability of the master versus beginner. This is to be expected if we postulate that experts in general are more efficient at searching their solution space.

Related to time to solution is another quantifiable difference--the pause times between retrieving successive equations or chunks of equations. Larkin (1979) has claimed that a number of physics equations are retrieved by the experts in succession, with very small interresponse intervals, followed by a longer pause. Her novice did not seem to exhibit this pattern of pause times in equation retrieval. This is interpreted to suggest that experts group their equations in chunks, so that the eliciting of one equation perhaps activates another related equation, so that it can be retrieved faster. (There is also some evidence that the chunk is associated with a "fundamental principle" of physics, such as Newton's Second Law, or Conservation of Energy.) Additional evidence for the rapidity of equation retrieval by the experts was demonstrated by Larkin (1981) when she found that experts were four times faster than the novices in accessing and applying equations during problem solving. This suggests to Larkin (1979) that for the experts, physics equations are stored in chunks or related configurations, so that accessing one principle leads to accessing another principle. This result is appealing because it is reminiscent of the chess results, where chess pieces were found to be chunked when the interpiece pause times during recall of a chess position were examined.

Another interesting aspect of novice problem solving is not only that they commit more errors than experts, but that even when they do solve a physics problem correctly, their approach is quite different. It is this difference that we want to understand, as well as why they commit errors. Likewise, it is also interesting to understand the circumstances under which experts make errors.

Qualitative Differences

Qualitative differences between an expert and novice problem solver are harder to define operationally, especially in empirical studies. However, it is the qualitative differences that distinguish expertise most noticeably. One prominent yet elusive difference between the expert and novice is that expert physicists, as noted before, seem to apply a "qualitative analysis" (Larkin, 1977a; Larkin, 1980; McDermott & Larkin, 1978) or "physical intuition" (Simon & Simon, 1978) to the problem, prior to the actual retrieval of physics equations. There are several possible interpretations of what constitutes "qualitative analysis." One interpretation is that "qualitative analysis," occurring usually in the beginning phase of problem solving, is the construction of a physical representation, that is, a representation that has some external, concrete physical referents. This ability to represent the problem physically in terms of real-world mechanisms was first noted over a decade ago, although not in the context of the expert-novice distinction. Paige and Simon (1966) observed that when algebra word problems that corresponded to

physically unrealizable situations were presented to subjects, a few of them immediately perceived the "incongruity" in the problem, whereas others proceeded to evoke equations before realizing that the solution was meaningless (such as a negative quantity for the length of a board). The former solvers apparently imagined the physical referents of the objects mentioned.

In physics problem solving, the construction of a physical representation may be helpful, or even necessary, for several reasons. First, Simon and Simon (1978) suggested that physical representation provides a basis for generating the physics equations. Second, physical representation provides a situation that can be used to check one's errors (Larkin, 1977a; Simon & Simon, 1978). Third, the physical representation provides a concise and global description of the problem and its important features. And finally, we conjecture that the physical representation permits direct inferences to be drawn about certain features and their relations that are not explicit in the problem statement, but can be deduced once a representation is constructed.

However, there is also reason to think that what occurs during "qualitative analysis" is more than the construction of a physical representation, since the often complex physical configuration and intuition deriving from what happens in a physical situation, may not necessarily lead to correct inferences. As the aforementioned work of Champagne and Klopfer (1980) and McCloskey et al. (1980) have indicated, naive problem solvers must not always rely on their physical intuition for constructing a representation. However, since

it is predominantly the experts who construct an elaborate representation, we postulate that this representation need not correspond directly to a physical representation, but may be more abstract.

A second qualitative difference between the expert and the novice, observed by Simon and Simon (1978), is in the number of "metastatements." "Metastatements" are comments made by the subjects about the problem-solving processes. On the average, their expert made only one metacomment per problem, whereas the novice made an average of five metacomments per problem. They were usually observations of errors made, comments on the physical meaning of an equation, statements of plans and intentions, self-evaluation, and so on.

There are several possible explanations for why their expert made fewer metacomments. First, he might be better at recognizing the correctness of a solution, so that he need not voice any uncertainties, etc. Secondly, their expert may have multiple ways to solve a problem (Simon & Simon, 1978), so that he can easily doublecheck his solution. Finally, the expert might have a well-structured representation of the problem to check his results against.

Another blatant qualitative difference between the solution processes of experts and novices lies in their solution paths (sequence and order of equations generated) (Simon & Simon, 1978). The important distinction between the expert and the novice is that the expert uses a "working forward" strategy, whereas the novice uses

a "working backward" strategy. The expert's strategy is simply to work from the variables given in the problem, successively generating the equations that can be solved from the given information. The novice, on the other hand, starts with an equation containing the unknown of the problem. If it contains a variable that is not among the givens, then the novice selects another equation to solve for it, and so on. (These processes and models based on them will be explained more fully later.)

This interpretation of the novice's performance initially seems counter-intuitive; that is, the novice's strategy appears to be more goal oriented and sophisticated. One interpretation of this difference is that the expert knows that he can achieve the goal simply by direct calculations of the unknowns from the givens. Another intepretation is that experts do not require complex planning for simple problems. They probably have existing routines or production systems that they can apply directly to the problems. This simple forward-working strategy of the expert does change, however, to a very sophisticated means-ends analysis of the goals and planning when the problems become harder (Larkin, 1977b).

A puzzling question concerning the difference between the two strategies is how people change from one to the other. Why is it that the expert can develop a more efficient system? One possibile answer is that over the years, the expert has built up and stored several fundamental sets of subroutines which can solve several types of basic problems. In this case, solving a problem becomes a matter of categorizing the problem into one or more problem types and applying

the existing subroutines. As we shall describe later, this ability to quickly categorize the problem is faciltitated by a powerful parsing mechanism that translates key words in the problem statement--words such as "at the moment," "catch-up," etc.--into problem types.

The second question is how can the expert construct a more efficient subroutine, if one does not already exist for solving a complex problem? We think that this facility lies in the rich internal representation that the expert has generated, a representation that permits many appropriate inferences to be drawn so that the problem can be simplified and reduced.

In sum, the analysis of the qualitative aspect of protocol data raises a number of important questions: Why is the initial "qualitative analysis" of the problem important? What kind of representation of a problem is constructed during this initial stage of analysis? Why are the sequences of equations generated by experts and novices different? What enables an expert to generate a sequence of equations that is more efficient? The quantitative analysis of the protocol data simply confirms a number of intuitions that we already have, but cannot explain: experts commit fewer errors, they can solve problems faster, and they seem to store related equations in closely knit chunk structures. Moreover, none of these quantitative findings provides any answers to the qualitative questions. Nor do they answer our questions posed earlier, namely, why are novices less successful at solving physics problems, and why are their procedures somewhat different, even when they are successful. Answering these questions is the focus of our own experimental program, which is described in

the latter part of this paper. These questions also drive current research and theory; we now turn to considering the current state of theory.

Theoretical Models of Physics Problem Solving

There has been a great deal more theoretical than empirical work done on problem solving in physics. In this section, we will review all those models that exist. They are of two types: psychological models that explicitly attempt to simulate human performance and artificial intelligence models which do not (although they may contain components that are simiar to human peformance). Both types of model are written in the form of computer programs.

Psychological Models

The majority of psychological models discussed here have several things in common. First, the behaviors they simulate are generally think-aloud protocols gathered while a person solves a physics problem. Second, except for one case, most of them solve mechanics problems taken from a first course in physics. Although these problems are straightforward, they are by no means simple. They do require some thought and usually take at least two minutes to solve. Third, the aspects of protocols that the models attempt to simulate are generally the sequences of equations generated by the solver. Hence, the qualitative aspects of the protocols (such as the initial analysis of the problem, the metastatements, and so on) are usually ignored. Finally, the simulation usually takes the form of a

production system.

To be more specific, the core of several of these models is a symbol-driven process. The variables representing the knowns and unknown(s) (the answer) in the problem are simply compared to the variables appearing in the various formulas that the model has in its possession. Two very simple selection criteria can be applied to produce two different behaviors. On the one hand, a formula can be selected in which all variables but one are knowns. That one unknown variable can then be asserted to be known (tagged as solvable, without any actual algebraic or arithmetic computation) and the process can be repeated until the new known is the answer to the problem. This is a working forward strategy typical of experts. On the other hand, a formula can be selected because it contains the desired unknown. If all the other variables in the formula are known then the problem is solved. If not, the unknown variable (the models discussed here generally discard a formula if it has two or more unknowns) becomes a new desired variable and the process is repeated; this is the working backward strategy characteristic of novices.

To make these two strategies more concrete, consider the following very simple example: There are two formulas available, one relating the variables a, b, and e and the other relating d, c, and e:

(1) $e = f(a,b)$

(2) $d = f(c,e)$

Suppose a problem is proposed such that a, b, and c are given (the knowns) and d is the desired answer (the unknown). The forward working method chooses equation 1 first, since a and b are known,

allowing the calculation of e. Since c and e are now both known, equation 2 can be selected and used to find d. By contrast, the working backward method chooses equation 2 first since it involves the desired unknown d. Since e is unknown, it becomes the intermediately desired unknown, and equation 1 is then chosen. Equation 1 can now be solved for e, which is substituted into equation 2 to find d.

Simon and Simon models. The first models to be discussed use the two strategies described above--working forward and backward. In the Simon and Simon (1978) models, the behaviors of two subjects, one novice and one expert, working a series of kinematics problems (describing motion in a straight line without any consideration for the causes of that motion), are simulated by two very simple production systems. The available formulas are represented in the conditions of the productions as lists of the variables they contain. The problem itself is presented as a list of the known and desired variables it contains. As explained above, the expert productions match the knowns in the problem with the independent variables in the formulas, while the novice productions match the desired unknown against the independent variable and the knowns against the dependent variables. The productions are listed in different orders, reflecting the fact that the two subjects sometimes used different formulas where both strategies might be expected to choose the same one. These two versions of the model simulate the equation selection behavior of the subjects quite well.

In this theory, there is no need to postulate any differences in the mechanism by which equations were produced; it is only necessary

to specify a difference in the order in which they were generated. Nor is skill difference attributable to trivial differences such as the lack of certain formulas. Both the expert and novice systems contain basically the same set of equations.

Knowledge development and means-ends models. Two related models are described in Larkin, McDermott, Simon, and Simon (1980). One is referred to as the Knowledge Development model, which simulates expert behavior, and the other is the Means-Ends model, simulating novice behavior. These models expand and improve on the Simon and Simon models in several ways to reflect more accurately human information processing capacities and the behavior of the subjects. Three separate memories are present: Long-term memory (LM), working (short term) memory (WM), and external memory (EM). LM consists of the productions themselves, which contain the necessary physics and procedural knowledge. WM is a small memory limited to about 20 elements and it is the contents of this memory that the condition sides of the productions are matched against. EM represents the pencil and paper used by a problem solver. The complete problem statement resides in this external memory and elements can be periodically transferred back and forth between EM and WM by the actions of certain productions to simulate the changing focus of attention of a problem solver and the process of recording intermediate results on paper.

The solution process begins with the problem statement in a coded form that specifies the objects involved, their attributes and points of contact, instants and intervals of time and the desired unknown(s).

(The complex problem of natural language understanding is avoided.) Both models have productions that assign variables to the necessary elements of the problem so that the appropriate formulas may be selected. As before, the two basic selection strategies, forward and backward, are employed but they are more elaborate to more closely simulate behavior.

The differences between the current and the previous Simon and Simon models are the most marked in the selection of a formula in the Means-Ends novice model, because novices are observed to do this in several discrete stages, first selecting a formula, then relating its variables to items in the problem, and then using it. A formula is originally selected for consideration if it merely contains a desired quantity. In cases where more than one formula contains the desired quantity, selectors tailored to represent observed novice preferences pick one. This model produces the same backward chain of equations as the earlier model. It then "solves" them by chaining forward, marking each previously unknown variable as known until the originally desired variable becomes "known." (Neither of these models has any actual algebraic manipulation ability.)

The Knowledge Development model is more similar to the previous Simon and Simon expert model. This is because experts generally do not exhibit the step by step behavior of stating an equation and then connecting it to variables in the problem. Thus, as before, the selectors select a formula on the basis of the unknowns and assert that the dependent variable is now known in one step. This situation can be viewed as a "collapsed" or over-learned version of the novice

model (this will become clearer shortly when other models are discussed). The main new feature of the model is that when more than one formula can be selected based on the knowns, information from the problem is used to decide among them. For instance if a (acceleration) and t (time) are knowns, then both $x=\frac{1}{2}at^2$ and $v=at$ could be selected. If the problem contains an object falling or rolling from rest, the first is selected; in all other instances the second is selected, corresponding to the observed expert preferences. It is in this sense that the knowledge about the problem is used.

In addition to the differences mentioned above, the Larkin et al. (1980) models have the ability to solve more kinds of problems than the previous ones, which were confined to kinematics. They solve dynamics problems (describing the motion of a body by considering the forces causing or influencing that motion), using two basic methods for solving such problems, Forces and Energies, and because they contain more than one solution method, they have an attention focusing mechanism. If a model is solving a problem using Energies, it should not try a Force equation halfway through the solution, nor should it select an equation when it is not through writing a previous one. To accomplish this focusing, goal elements are included in the conditions of many of the productions. At the beginning of a solution process, a goal is set (placed in WM and EM) so that only productions related to that goal can execute.

Able models. The Able models of Larkin (1981) address a different issue than strictly simultating the problem solving processes. Instead, they attempt to simulate the learning processes,

that is, how a novice might become an expert. In the model's "naive" state, it is called the Barely Able model, and after substantial learning, it is called More Able. The learning process is modeled by a mechanism for adding procedures that is generally used in adaptive production systems (Waterman, 1975).

Barely Able starts with a list of equations that can be used in the Forces or Energy methods, and operates with a general means-ends strategy for applying them that is similar to the previous Means-Ends model. The learning process itself is quite straightforward: Whenever a production succeeds in applying an equation to derive a new known value, it creates a new production that has the previous knowns on the condition side and an assertion of the new known on the action side. For example, if Barely Able solves the equation $V=V_0 + at$ for a, then the new production will check to see if V_0, V and t are known and, if so, assert that a is known. Psychologically, this means that the procedure for finding the right equation and solving for the unknown becomes automated once the initial production has been executed. Thus, as Able solves more and more problems, it looks more and more like the Knowledge Development model mentioned earlier--it becomes forward-working because all the backward-working steps become automated.

There are two limitations to the Able model. The first is that the learning takes place in one trial. This is psychologically unrealistic and a more complicated learning function probably needs to be built in which some aspects of learning take place faster than others. The second limitation is that the model does not provide the

capability to concatenate series of productions into one, (Neves & Anderson, 1981). Such a mechanism would allow two or more formulas to be combined into a single step, as experts are often observed to do.

Model PH632. A model labeled PH632 developed by McDermott and Larkin (1978), has a somewhat different focus than those previously described. Its purpose is to examine and model in a general way the use of problem representations by an expert solver, but not to exhibit a detailed psychological model of the process. It is, again, a production system with external, working, and long term memories. The condition sides of the productions can contain goal elements that keep attention focused on the specific task at hand and that allow the productions to be hierarchically organized.

A series of four representational stages of a problem is postulated: verbal, naive, scientific, and mathematical (see also Larkin, 1980). The model assumes that a problem solver progresses through these stages as a problem is solved. However, the detailed description of the model (McDermott & Larkin, 1978) starts with the naive representation. The naive representation is a sketch depicting the components of the problem and their relationships, and is implemented as a data structure that encodes this information. The scientific representation contains abstract physics concepts such as forces, momenta and energies (which must generally be inferred by the problem solver) and is usually depicted as a free-body diagram. The mathematical representation consists of the equations relating the variables in the problem that must be solved to produce the final answer.

Once PH632 has a naive representation, it tries one of the two solution methods mentioned earlier--Forces and Energies. If both are adequate, the one chosen may simply be the first one tried. Once a particular method is chosen, its productions give the model the abiity to scan the sketch qualitatively to determine where the objects and systems of interest are, whether they are familiar or unfamiliar, and how they are related. If a system is familiar (such as a hanging block), PH632 can use its knowledge to build a production describing it. If the system is unfamiliar, an extended analysis is conducted to produce an encoded version of a free-body diagam. This difference in representation corresponds to an expert's tendency not to draw an explicit free-body diagram of a familiar system. The model makes qualitative checks as it proceeds to determine whether its representation seems correct and whether its approach is working. For instance, in a statics problem (one with no motion), it checks to make sure all of the forces are balanced by at least one opposing force. It can also test whether all of the entities generated in the scientific representation, such as forces, can be related to the quantities given in the problem statement so the equations can be generated.

Once assurance is gained that the model is on the right track, it can write the equations for the mathematical representation. Because all of the forces have already been located and resolved into components in construction of the scientific representation, this step is relatively simple. Unlike the previous models, PH632 can perform

the algebraic and arithmetic operations necessary to produce the answer.

<u>Atwood</u>. Larkin's (1980) latest program, Atwood, concentrates on the verbal representation stage, an area generally ignored by the previous models. Considering the difficulties and complexities encountered by AI researchers in building language understanders, Atwood accomplishes its task in a surprisingly simple and straightforward way. Because mechanics problems in general contain a rather small set of basic objects, attributes, and relationships, it can simply ignore most of the words in a typical problem statement and concentrate on the key words.

Basically, Atwood contains a set of schemas that tell it what words to attend to and what situations those words may indicate. Thus, it knows that the word "rod" is important and that there should be one and only one length associated with it. "Pulley" is another keyword and Atwood's schema tells it that there will be a rope passed over this object and that the rope should have objects connected to each end.

Using some rudimentary knowledge of English syntax, Atwood processes the problem statement word by word, creating nodes for each physics object it recognizes and connecting these nodes into a semantic net with the help of the knowledge of their legal relationships contained in the schemas. When tested on a set of 22 of the problems collected by Chi, Feltovich, and Glaser (in press), Atwood was able to build correct nets for 15 of them, while ignoring roughly two-third of the words they contain.

Summary and discussion of the psychological models. The psychological models so far developed, focus their attention on the different approaches that experts and novices take in terms of the sequence of equations they generate--forward working versus backward working. In these models, it is assumed that experts are forward working because their initial backward solution procedure becomes automated with learning. The question of initial problem representation is generally avoided in these models, perhaps primarily because it is difficult to obtain empirical information on this process solely through the usual forms of protocol analysis. As we shall describe later, other techniques are required for this purpose.

An alternative theoretical framework is to suggest that novices are data driven. They treat the unknown and known variables as literal symbols and plug them into equations in their repertoire. Experts, on the other hand, are schema driven in the sense that their representation of a probem accesses a repertoire of solution methods. Hence, for the expert, solving a problem begins with the identification of the right solution schema, and then the exact solution procedure involves instantiation of the relevant pieces of information as specified in the schema. This is particularly likely because mechanics problems are overlearned for the experts, especially experts who have spent a great deal of their time teaching. Another interpretation is to postulate that novices also solve problems in a schema-driven way, except that their schemas of problem types are more incomplete, incoherent, and at a level hierarchically lower than those possessed by the experts. In our opinion, the development of

psychological models should proceed in this particular direction, building knowledge structures in the forms of schemas, in order to capture the problem-solving processes of experts and novices. Some empirical evidence for the validity of this interpretation will be presented later.

Artificial Intelligence (AI) Models

AI programs, unlike those previously discussed, are not specifically intended to model observed behavior or to take into account theories of human cognitive architecture. Their general aim is to successfully solve physics problems by any means possible. However, they do contain elements that are very similar to both human behavior and the previous psychological models.

One of the main issues addressed by the AI models is representation--how to represent the knowledge the program needs to form a representation of the problem and solve it. Indeed, the current recognition in psychology of the importance of representation probably derives from the early recognition of its importance in AI and computer science in general. The question of how physics knowledge is represented is a major research problem, as the rudimentary state of such representations in the psychological models indicates.

The first phase of a problem solution is reading and understanding (or translating) the verbal problem statement. Much work has been done on the general problem of natural language understanding in AI and two of the programs to be desribed put

considerable emphasis on this stage. Both are more detailed and complex than the simple Atwood (Larkin, 1980) translator, since they aim for a complete translation utilizing all of the information in the problem statement. Thus, both use esoteric translation processes and have extensive knowledge bases of syntactic and semantic information, including specific physics knowledge in a well-organized form to allow a correct physical interpretation of a problem. Once translation is complete, some kind of language-free, internal computer model of the problem exists, which can be compared to a naive representation.

Issac. Issac by Gordon Novak (1977) is a program that can read the problem statement. It does this for statics problems only. The key feature of interest is the representation of objects as idealized physics entities. For instance, in a problem that has a person *standing on a* ladder, the properties that are important to the solution are his mass and location on the ladder. He can therefore be represented as a "point mass." But if he is holding up one end of the ladder, only the point on the ladder he is holding is important and he becomes a "pivot." This idealization is accomplished in Issac by using Canonical Object Frames (schemas) from the knowledge base. Each one contains the knowledge necessary to abstract the proper characteristics from the "real life" object and to use the idealized object properly in the solution of the problem. This idealization process corresponds only partially to the formation of scientific representation because no attempt is made to represent or analyze qualitatively the other essential physics entities in a statics problem--the forces. Instead, all possible balance-of-forces

equations are written at each point of contact between objects, resulting in many more equations than are actually needed for a solution. This illustrates the problems that can arise if the representation of a problem does not generate an efficient solution.

Newton. Newton by Johann de Kleer (1977) does not have any language translation facility. It solves roller coaster problems (blocks sliding on curved surfaces), and they are best represented as a picture of the track, which is provided in a symbolic form. The key feature of this program is a process of qualitative analysis referred to as envisionment. Newton envisions, as a human solver might, what might happen to the sliding block based only on the general shape of the track. Thus, on an upslope the block might slow down and slide back down or continue up. At the crest of a hill, the block might be traveling so fast that it flies off into space or it might slide down the other side. Using a series of production rules that codify such qualitative knowledge, Newton builds a tree of possible paths of the block that guides further processing of the problem. Some simple problems may be solved using only this qualitative reasoning. If this is not possible, then schemas are used that contain knowledge and formulas necessary to analyze each node of the tree (section of the track) mathematically. In cases where the value of a particular variable is needed for the answer, the familiar means-ends process is used to choose the proper formulas to apply.

Mecho. Another language translator is Mecho by Bundy, Byrd, Luger, Mellish, & Palmer (1979). This program solves problems from kinematics and those with pulleys. It has also been extended (Bundy,

1978; Byrd & Borning, 1980) without translation to solve problems in statics and roller coasters in an attempt to make the problem-solving part as general as possible by encompassing the work of others (e.g., de Kleer, 1977; Larkin & McDermott, 1978; Novak, 1977). The salient feature of this program and, perhaps, the key to its extendability, is a two-level knowledge organization. On the object (lower) level is the physics knowledge, organized as rules and schemas and the problem itself. The problem passes through several stages of representation on the way to a solution. For example, the natural language translation feature produces a symbolic representation specifying the objects in the problem and their properties. Where necessary, schemas describing important objects, such as a pulley, are cued in from the knowledge base. Thus, this initial internal representation might be viewed as naive with elements of a scientific representation. The next general step is to produce the mathematical representation, which can then be solved algebraically. This is not a simple step however. The meta-(upper)level of the knowledge base contains all of the procedural knowledge necessary for the entire solution process, organized as a set of rules and schemas. It includes rules for interpreting the object level knowledge for use at each step of the process, for making inferences when needed information is not explicitly stated, for deciding upon a general solution strategy, for selecting equations (means-ends strategy again), etc. Athough a complete scientific representation is not explicitly formed, the planning and inferencing powers of the meta-level implicitly use the elements of such a representation to plan the solution before

equations are actually generated. Thus, in a statics problem, for instance, the planning process eliminates the problem of excess numbers of equations experienced by Issac.

The organization of procedural knowledge into explicit modular form is what is most interesting psychologically about Mecho. Quite often, such knowledge is buried in the structure of a program and the assumptions that went into writing it, making changes difficult and modeling of procedural learning impossible. This two-level organization also allows the declarative knowledge to be present in only one form, which can be interpreted by the meta-level for use at each step of the solution process. By contrast, both Issac and Newton contain separate representations of the same physics knowledge for each step. In a sense, Mecho can learn (though not on its own) and has learned to solve new problems in a fairly realistic way psychologically because all that is necessary is to give it other new pieces of procedural and declarative knowledge.

Summary. Although as noted, the purpose of these AI programs is not to model human behavior, it is clear that they contain many psychologically important features and ideas. The question of representation of the problem and the knowledge base is common to both fields and the proposed solutions--stages of representation, rules, schemas, (often called frames in AI)--are generally similar. However, since AI is not limited by empirical knowledge of behaviors, these programs can venture into areas where psychological model builders have more difficulty simulating, such as natural language translation, qualitative analysis (e.g., envisionment), planning and inferencing

processes, and the actual specification of knowledge organization. The importance of these items to the success of AI progams emphasizes the need for much more work to determine empirically how they occur in humans.

EMPIRICAL STUDIES TOWARD A THEORY OF EXPERTISE

The objective of the series of investigations that we have carried out is to construct a theory of expertise based upon empirical description of expert problem-solving abilities in complex knowledge domains. In this case, the knowledge domain is physics, in particular, mechanics. There are basically three questions that guide our efforts. First, how does task performance differ between experts and the novices? This question has been partially answered in the review of empirical evidence on physics problem solving. To recapitulate, the basic differences found thus far are: (a) the two groups use different strategies for solving problems, forward versus backward; (b) they seem to have different chunking of equations; (c) in an initial phase of problem solving, experts tend to carry out a "qualitative analysis" of the problem; and (d) experts are faster at solving problems. One of our goals is to describe more extensively these differences between experts and novices.

The second question asks: How are the knowledge bases of skilled and less skilled individuals differently structured? It is clear that the skilled individual possesses more knowledge, but how is that knowledge organized? Again, some research has already addressed this issue. Simon and Simon (1978) initially postulated a difference in

the knowledge base in terms of the conditions of the productions. Larkin (1979) has postulated a difference in the way equations are stored. Experts store them in relation to a high level principle, but this this does not seem to be the case for the novices. In our work and in Larkin's model Atwood (1980), knowledge is postulated to be organized in the forms of schemas.

The third question guiding our work is: How does the organization of the knowledge base contribute to the performance observed in experts and novices? The relation between the structure of the knowledge base and solution processes must be mediated through the quality of the representation of the problem.

A problem representation, as we stated in Chi, Feltovich, and Glaser (in press) "is a cognitive structure corresponding to a problem, which is constructed by a solver on the basis of his domain-related knowledge and its organization." We adopt Greeno's (Riley, Greeno, & Heller, 1981) notion of a representation, which takes "the form of a semantic network structure, consisting of elements and relations between these elements" (p. 23). Hence, we hypothesize that at the initial stage of problem analysis, the problem solver attempts to "understand" the problem (Greeno, 1977), i.e., constructing a representational network containing elements specifying the initial state of the problem, the desired goal, the legal problem solving operators, and their relational structures. From such a structure, new inferences can be deduced. Hence, the quality, completeness, and coherence of an internal representation must necessarily determine the extent and accuracy of derived inferences,

which in turn may determine the ease of arriving at a solution and its accuracy. Therefore, the quality of a problem representation is determined not only by the knowledge that is available to the solver, but the particular way the knowledge is organized. One way to capture empirically the difference between the representation of the expert and that of the novice has been the amount of "qualitative analysis" occurring in the beginning of the problem solving processes.

Because of its apparent overriding influence on problem solution (Hayes & Simon, 1976; Newell & Simon, 1972), we have focused our studies mainly on the representation of a problem. We employ methods of tapping knowledge in ways other than the analyses of problem solving protocols, since as we will see shortly, the analyses of protocols often provide limited information. However, the first study we describe examines the protocols of problem solving, to see what kind of information they do provide, as well as to see in what ways they provide a limited glimpse into the knowledge structure. The next set of studies looks at the categorization behavior of problem solvers; the third set of studies looks at the knowledge available to individuals of different skill levels; and finally, the fourth set of studies examines the features in a problem statement that might elicit the categorization processes--or to put it another way, what is considered to be the relevant features of a problem by experts and novices.

Study One: Protocols of Problem Solving

In this first study, we attempted to characterize and contrast, both quantitatively and qualitatively, the problem solving processes of experts and novices, beginning with the reading of the problem, through to the checking of the solution. To do so, the problem solving protocols of two experts and two novices solving five mechanic problems were examined. This study was initiated and carried out by Joan Fogarty. The specific goals were twofold: First, we wanted to describe some quantitative parameters of expert and novice problem-solving processes, and compare these data with those existing in the literature; second, we wanted to contrast some qualitative differences between experts and novices, particularly focusing on the qualitative aspects of the analyses of the problem.

The five mechanics problems used in this study were taken from Chapter 5 of Halliday and Resnick (1974). The expert subjects for this study were two professors of physics who had considerable experience teaching introductory physics. The novices were two freshmen physics majors (A students), who had just completed a term of undergraduate physics, using Halliday and Resnick (1974) as the textbook, in which mechanics problems of the type used in this study were taught. Each subject was presented with written problems, one at a time, and was instructed to "think aloud" while he solved the problems.

Quantitative Results and Discussion

A variety of quantitative measures can be obtained from protocol data. These are elaborated below.

Errors. The experts on the average, made 1 out of 5 posible errors, whereas the novices made three out of five errors. (See Table 1. Errors are marked by parentheses around the solution time. If only a part of a problem is incorrect, then that part is indicated by a subscript.) As anticipated, experts made fewer errors than novices. The fact that one of the experts made two errors suggests that these problems are nontrivial. On the other hand, these are problems that a competent novice (A student) can solve. Novice K.W., for example, solved 4.5 out of the 5 problems correctly.

Solution times. Solution times were determined by timing the length of the protocols. Looking only at the correct solution times for the entire problem (see Table 1), the mean solution time for the experts averaged about 8.96 minutes, whereas the average correct solution time for the novices was 4.16 minutes. The magnitude of our solution time for problem solving protocols is much longer than those obtained by Simon and Simon (1978). Their problems were selected from a high school physics text and were limited to kinematics; such problems can be solved mainly through algebraic manipulation. Our problems were more complex; they were chosen from a college physics text and involved dynamics, which requires that forces be explicitly taken into account. Applying The Force Law requires some physical inferences to be made before equations can be brought into play.

The novices in this study actually solved problems faster than the experts. However, this seems to be an artifact of the great number of errors made by Novice C.H. That is, Novice C.H.'s only correct solution was problem 1, which in fact, took him longer than the rest of the subjects to solve. But, because problem 1 happens to be a short problem, and since that was the only problem he solved correctly, his average latency was reduced, because it was determined by the speed of solving that particular problem. Novice K.W.'s. solution times, on the other hand, are actually comparable (averaging 7.01 minutes) to the experts' (averaging 8.96 minutes).

The only obvious outlier in solution time occurs in problem 2, where Expert R.E. took significantly longer than Novice K.W. Examining the protocols in detail, we see that Expert R.E. in this case sought and calculated a value unnecessarily. When he discovered that the problem was really much simpler than he thought, the actual protocol for the short solution took only about 1.33 minutes.

Hence, barring unusual circumstances, competent novices not only can solve these problems, but they can do so in approximately the same amount of time as experts. However, if the task had emphasized speed, the experts probably could have solved the problems much faster than the novices. We suggest, however, that protocol data are not a particularly viable way to assess the speed of problem solving.

Number of quantitative relations. Another quantitative parameter that may shed some light on skill differences between experts and novices is the number of quantitative relations generated by the subjects as they solve problems. Table 1 also shows the total number

of quantitative relations generated by each subject for each problem. A quantitative relation is defined as any mathematical relation among physical entities, and it generally takes the form of an equation. Excluded were algebraic manipulations of already generated equations, and instantiations of equations (i.e., substituting values for the variables). In general, there appear to be no systematic differences in the number of quantitative equations generated as a function of skill. There was greater variability in the number of equations generated by a given subject for the different problems, than between subjects on the same problem.

"Chunks" of equations. As stated earlier, Larkin (1979) has hypothesized that experts store physics equations in tightly connected "chunks," whereas novices store them individually. To test the "chunking" hypothesis, Larkin (1979) measured the times during the problem solving process when quantitative relations were generated. Her results showed that the expert generated a great many pairs of equations with short pauses between the equations, whereas her novice generated fewer equations with shorter pauses.

Using the same analysis, we also examined the distribution of generated equations over time. For each subject, the time interval between the generation of each pair of quantitative relations was calculated for each problem. Our data do not discriminate between the generation pattern of experts and novices. If anything, the results indicated that the opposite was true. That is, the novices seemed to have generated a greater number of relations in close succession.

There are substantial individual differences, however. Novice C.H. showed the strongest degree of "chunking," or generated the largest number of quantitative relations in rapid "bursts." How do we account for the discrepancy between our results and Larkin's? One interpretation is to hypothesize that a burst of equation generation may be an artifact of various problem solving strategies that subjects may adopt. Our novice subjects, for example, reported that when they get stuck on a problem, they generate as many related equations as they can think of on paper. They then look at the equations they have generated to get some hints about how to proceed. This would produce clusters of equations.

Another strategy, reflecting the style of solution processes of individual subjects, relates to the way equations are generated, that is, often all at the same time. Novice C.H., for example, would spend a considerable amount of time generating equations. This pattern of solution processes would necessarily inflate the number of equations generated within a short period of time. Perhaps the generating of equations in bursts may also be the outcome of another artifact, discussed in the next section: the drawing of free-body diagrams.

Even though we did not replicate Larkin's (1979) finding that experts tend to generate equations in clusters, this does not deny the possibility that the storage of equations may indeed be different in the knowledge base of the experts and novices. Our conclusion is that protocol analysis of equation generation will not address this particular issue directly. In order to address the issue of how equations are stored in the knowledge bases of experts and novices,

one needs to design a study where experts and novices are asked to generate or freely associate equations outside the context of a problem solving situation.

Number of diagrams generated. Another potentially interesting quantitative measure is the number of free-body diagrams drawn by the subjects. The construction of free-body diagrams, appears to form an important component of problem solving. Free-body diagrams are partial figures that depict partial abstractions of the total physical situation. They may be drawn for all or part of the physical situation, and utilize directional arrows denoting the forces acting in a physical system.

The number of diagrams including free-body diagrams drawn by each subject for each problem is also shown in Table 1. Again, there appears to be no systematic skill differences, although there seems to be some individual differences, with Expert R.E. and Novice C.H. drawing the greatest number of free-body diagrams. These two individuals also generated the greatest number of equations, and also produced the greatest amount of clustering.

Drawing free-body diagrams may inflate the number of equations generated in clusters. Both novices, as well as the experts to a lesser extent, utilized the strategy of constructing free-body diagrams, which is taught and emphasized in introductory physics courses. Using the free-body diagrams, equations relating the forces can be generated. Hence, the more frequently a subject draws a free-body diagram, the more likely he is to have clusters of equation generation. Therefore, bursts of equation generation may be an artifact of a solver's need to generate many diagrams.

It is not clear to us what the purpose is of generating many free-body diagrams. We speculate that when a problem is difficult for a subject, the subject tends to draw more diagrams. Each drawing may be seen as an attempt to create a meaningful representation of the problem. For example, for problems that took the longest to solve, a large number of diagrams tended to be generated (such as problem 2 for Expert R.E.). Furthermore, problem 2 was the one that Expert R.E. had some difficulty with, having derived a value unnecessarily. Likewise, for Novice C.H., problem 3 took the longest time to solve (which he did incorrectly); he also generated the greatest number of diagrams for that problem. These speculations need to be confirmed, but it seems that drawing free-body diagrams may be a way of helping the subject to create a meaningful representation. It may also indicate that the subject is having difficulty going beyond the visual stage of problem representation.

In another study (Study Five in this paper), when four experts and four novices were asked to solve a problem, the novices generated four times as many (4.7) diagrams as the experts (1.0 diagrams). The novices had more difficulty solving the problem correctly (3 out of 4 errors) than the experts (1 out of 4 errors). This provides some additional support for the notion that frequent generation of diagrams is used as an external aid to create a meaningful problem representation, and especially when subjects are having difficulties.

Summary of quantitative measures. The results of this study indicate that few of the quantitative measures we used meaningfully

differentiated the experts from the novices. The quantitative measures obtained from protocols seem to be tenuous measures that are confounded with individual differences and the particular strategies adopted by the problem solver. We now turn to qualitative analyses of the protocols to locate differences that can be attributed to skill.

Qualitative Results and Discussion

For reasons already indicated, and since a great deal of attention had been devoted to the equation generation and manipulation stages of problem solving, in this section of the data anlayses, we will focus our attention on the initial "qualitative analysis" stage of problem solving. We assume that during this stage of processing, a representation of the problem is constructed, and that this occurs primarily during reading of the problem, and is completed in the first 30-40 seconds after the problem has been read. We estimate that this stage takes a very short time because it appears to be analogous to the stage of "initial analytical assessment" that Simon and Barenfield (1969) talked about for chess problem solving, and the stage of "preconception" that expert musical sight readers engage in prior to the actual playing of a musical piece (Wolf, 1976). The short duration of these initial processes is an important consideration in determining our subsequent experimental procedure.

Figures 2 [illegible] show two samples of protocols, one from Expert R.E. and the other from Novice C.H., both on the first part of problem 5. The protocols have been segmented into four types of episodes: "qualitative analysis," drawing diagrams (which may be

Expert R. E.
(Problem #5)

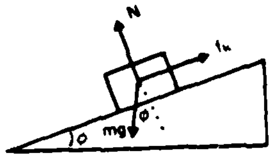
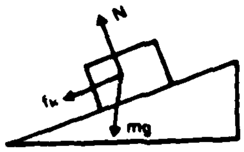
Taxonomy of Episodes	Physics	Protocols
*Qualitative Analysis (inferences)	Constant velocity ⟶ Frictional force Frictional force opposes force due to weight of block *Friction ⟶ Coefficient of friction $\sim$ angle ϕ	"There must be a frictional force retarding the motion because otherwise the block would accelerate down the plane under the action of its own weight...the angle ϕ must be related to the coefficient of friction somehow."
Drawing Free Body Diagram		"You would have a normal force perpendicular to the plane, the weight down, and the force of kinetic friction would lie along the plane...the angle between the weight vector and the normal to the plane is also angle ϕ."
Generate Equations	$mg\sin\phi - f_k = 0$ $N - mg\cos\phi = 0$ $f_k = \mu_k N = \mu_k mg\cos\phi$	"For motion down the plane would be mg times sinϕ minus f which is retarding things and that's equal to zero. For motion perpendicular to the plane, you would have the normal force acting upward, but mgcosϕ acting downward or into the plane and those two things sum to zero. The only relation you need in addition is that the force of kinetic friction is μ times the normal and is therefore μ times mgcosϕ."
Albegraic Manipulation	$mg\sin\phi - \mu mg\cos\phi = 0$ $\mu_k = \tan\phi$	"So substituting that ($f = \mu mg\cos\phi$) into the first equation, which I've circled, you would then have mgsinϕ, f which would be μ times mgcosϕ, and all of that would be equal to zero, and so what one finds then is that μ, the coefficient of friction must be tanϕ."
Reread Question A		

Figure 2. Expert R.E.'s protocol on problem 5, segmented into episodes.

Expert R. E.
(Problem #5) continued

Taxonomy of Episodes	Physics	Protocols
Draw Free Body Diagram		"So let's draw the plane again... the difference is that the frictional force...acts in the other direction."
Qualitative Analysis		"We know the initial speed is Vo...I'm sort of fishing here for a minute, the final speed... is obviously zero."
Generate	$V_f^2 - V_o^2 = 2ax$	"We have an expression which relates several things of interest to us...all at the same time."
Qualitative Analysis		"We can easily solve for x providing we know the other things in the equation....We don't know a but that's not hard to find."
Generate	$mg\sin\phi + \mu_k mg\cos\phi = ma$	"This time both mgsinϕ and the frictional force...those two forces act in the same direction."
Manipulate	$\cancel{m}g\sin\phi + \mu_k\cancel{m}g\cos\phi = \cancel{m}a$ $\mu_k = \tan\phi = \frac{\sin\phi}{\cos\phi}$ $a = g\sin\phi \times \frac{\sin\phi}{\cos\phi} g\cos\phi$ $= 2g\sin\phi$	"The masses cancel everywhere ...we also know μ_k...μ_k is the tangent of ϕ... which is the sin of ϕ over the cos of ϕ...the cosϕ's cancel and you're left with the acceleration down the plane of...twice gsinθ.
Qualitative Analysis (Inference) (check answer)	block slides uniformly $\longrightarrow f_k = F_g$ $= mg\sin\phi$ f_k now in opposite directions $\rightarrow$ Total Force $= F_g + f_k = 2mg\sin\phi$ $a = \frac{F_{total}}{m} = 2g\sin\phi$	"So effectively you have...an acceleration...of twice the weight... I n the first part of the problem...friction... must be exactly equal to gsinϕ and if you have it operating in the opposite direction..."

EXPERT R. E.
(PROBLEM #5) CONTINUED

TAXONOMY OF EPISODES	PHYSICS	PROTOCOLS
MANIPULATE	$0 - V_o^2 = 2(-2g\sin\phi)x$ $x = V_o^2/4g\sin\phi$	"Now let's go ahead and solve for...V Final squared was 0. V initial squared was what it is...so what you end up with for, for x is Vo squared over $4g\sin\phi$."

NOVICE C. H.
(PROBLEM #5)

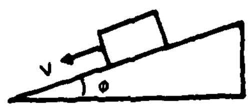
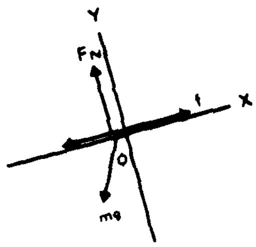
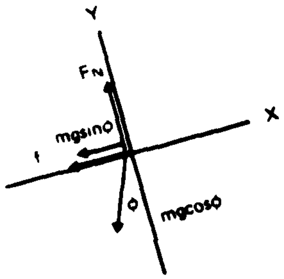
TAXONOMY OF EPISODES	PHYSICS	PROTOCOLS
DRAW DIAGRAM		"Let me draw a picture. An inclined plane with slop angle φ ...and it's (the block) sliding down the plane with a velocity ...constant velocity."
QUALITATIVE ANALYSIS (INFERENCES)	Constant velocity ⟶ $\Sigma F = o$ ⟶ friction	"Since it's (the block) sliding down the plane with constant velocity, it means the sum of the forces is zilch so there's a, there's got to be some kind of friction on the thing..."
DRAW FREE BODY DIAGRAM		"I'll draw a free body diagram. There's the weight mg, there's the frictional force, then there's the normal force perpendicular to the plane.
GENERATE EQUATIONS	Force parallel to plane = mgsinφ $F_N = mg\cos\phi$ $f = \mu F_N$	"Ok. So I'm going to draw trusty axes and resolve weight into a, into....You've got : there so this mgcosφ, and this is mgsinφ...normal force is going to be equal to mgcosφ and friction equals, umm...μ times the normal force."
MANIPULATE	$f = \mu mg\cos\phi$	"So that frictional force is equal to μmgcosφ."
GENERATE	$V^2 = V_o^2 + 2a(x-x_o)$	"The block is projected up the plane with an initial velocity. So I'm going to use...equation for motion $V^2 = V_o^2$ + 2 times acceleration times change in distance."

Figure 3. Novice C.H.'s protocol on problem 5, segmented into episodes.

Novice C. H.
(Problem #5) continued

Taxonomy of Episodes	Physics	Protocols
Manipulate	$x_o = 0 \quad V = 0$ $\frac{V_o^2}{2a} = x$	"Initial position I'm going to call 0...final velocity equals 0 so I get V_o(sic) over 2a is going to equal the x."
*Qualitative Analysis (Inference)		"a is going to be acceleration due to the frictional force." (WRONG)
Draw Free Body Diagram		"Now we've got a different drawing. We've got mg and the velocity is up the plane so frictional force...is down the plane."
Generate	$\Sigma F_x = ma$	"...sum of the forces in my x direction is going to equal mass times acceleration."
Manipulate	$mg\sin\phi + f = ma$ $mg\sin\phi + \mu mg\cos = ma$ $a = g(\sin\phi + \mu\cos\phi)$ $x = \frac{V_o^2}{(g\sin\phi + \mu\cos\phi)}$	"So, you've got mgsinφ + frictional force equals the mass times acceleration, so frictional force is equal to... μ times the normal force... my m's go out so the acceleration equals g times sinφ + μcosφ. So I substitute back in the other equation." (Leaves out factor of 2)

either the diagrams depicting the main components of the problem, or the abstracted free-body diagram), generating equations, and manipulating equations. There are several general remarks that can be made about the initial stage of the protocols.

Before proceeding with the discussion of the protocol data, it may be necessary to clarify a few terms and operational definitions. Any statements in the protocols that do not relate to drawing diagrams, generating and manipulating equations, were considered to be "qualitative analysis" of the problem. These statements can further be of a variety of types, such as references to planning, checking of the solution, and so on. We focused specifically on those "qualitative analysis" statements that seemed to generate knowledge not explicitly stated in the problem, that is, inferences. (These "qualitative analysis" statements are not to be confused with qualitative analysis of the protocol data.)

First, contrary to the picture painted earlier, the protocol data indicate that our novices also spent time analyzing the problem qualitatively. During this stage, some inferences about the problem are drawn. A simple count of the number of propositions that were made that can be judged to be inferences shows that experts make, on the average 12.75 propositions and novices make 10.58, which is not reliably different. Consistent with our earlier assertion, the initial episode of "qualitative analysis" is usually short in duration, taking only one paragraph in the protocols.

The second observation is that, unlike what is commonly believed, the "qualitative analysis" episode often occurs throughout the

protocols, not just at the beginning. For example, the inference episode occurs, on the average, 2 1/2 times throughout each problem for the experts and 1 1/2 times for the novices, although this difference is again not significant. Because of this phenomenon, it is difficult to ascertain exactly when the construction of a representation is completed. These protocols lead us to think that a gross representation is initially constructed; then if it needs to be refined, that can occur later in the protocol.

The third observation is that errors in solution have two sources. One source is trivial computation error, resulting either from faulty manipulation or instantiation of equations. An example of a trivial computation error occurs in the last episode of Figure 3. In manipulating the equations, the novice made an error by a factor of two. The other source of solution errors can be traced to either the generation of wrong inferences, or the failure to generate the right inference. The inference episode in Figure 3 having an asterisk beside it, indicates an example of a wrong inference. We attribute the source of solution errors in general to these incorrect inferences, even though in this particular case, this incorrect inference was not the cause for the problem's incorrect solution. This is because the novice was able to generate all the correct equations. The mistake in this problem arises from the solver's failure to complete the solution by substituting for μ. Incorrect inferences are relatively easy to detect in the protocols. What is more difficult to capture in these protocols, is the solver's failure to generate a necessary inference. This can be captured only by

comparing and contrasting the expert's and the novice's protocols, in trying to understand a novice's error. Our interpretation is that Novice C.H. did not complete the solution (see the last episode of Figure 3) because he failed to generate the inference that the coefficient of friction μ is somehow related to the angle ϕ, as did the expert (see Figure 2, the first episode). Without setting an explicit goal to relate the two (μ and angle ϕ), Novice C.H. could not solve the problem, even though he had all the necessary equations.

Hence, in general, we would conclude from examination of the inference generating episodes of the protocols, that both experts and novices are just as likely to spend time generating tacit knowledge about a problem, and both groups are just as likely to do so iteratively across the entire problem solving protocols. However, it is the quality of the inferences that matters. Novices are more likely to either generate the wrong inference, or fail to generate the necessary inferences. A large number of the novices' errors can be traced to this source.

Studies on the Categorization of Problems

To say that novices either fail to make the appropriate inferences during qualitative analyses, or that they do not generate inferences at all, does not explain the source of incomplete or erroneous inference making. To uncover this limitation of the novices, we have to understand the knowledge structure of the experts and novices, and how that knowledge enhances or limits their problem solving abilities. Analyzing the protocols of problem solving does

not appear to provide enough information of this kind. Our research described here, therefore, is concerned with ways of exploring the knowledge of a problem solver, through means other than analyzing solution protocols.

We hypothesize that a problem representation is constructed in the context of the knowledge available for a particular type of problem. Further, we make the assumption that the knowledge useful for a particular problem is indexed when a given physics problem is categorized as a specific type. Therefore, expert-novice differences may be related to poorly formed, incomplete, or nonexistent problem categories. Given this hypothesis, we investigated knowledge contained in problem categories. Our first order of business then, was to determine whether our initial hypothesis is true: that is, are there reliable categorizes to which problems are typed, and if so, are these categories different for novices and experts?

Evidence already exists to suggest that solvers represent problems by category, and that these categories might direct problem solving. For instance, Hinsely, Hayes, and Simon's (1978) study, found that college students can categorize algebra word problems into types, and that this categorization occurs very quickly, sometimes even after reading just the first phrase of the problem statement. This ability suggests that "problem schemata" exist and can be viewed as interrelated sets of knowledge that unify superficially disparate problems by some underlying features. We refer to the knowledge associated with a category as a schema. The chess findings of Chase and Simon (1973a, 1973b) can also be interpreted as showing that

choosing a chess move results from a direct association between move sequences and a chunked representation of highly stereotyped (or overlearned) chess pieces or patterns. There is also evidence in studies of medical diagnosis that expert diagnosticians represent particular cases of disease by general categories, and these categories facilitate the formation of hypotheses during diagnostic problem solving (Pople, 1977; Wortman, 1972).

Study Two: Sorting Problems

To determine the kinds of categories subjects of different experience impose on problems, we asked eight advanced Ph.D. students from the physics department (experts), and eight undergraduates (novices) who had a semester of mechanics, to categorize 24 problems selected from eight chapters (5 through 12) of Halliday and Resnick's (1974) Fundamentals of Physics. The subjects' task was simply to sort the problems based on similarities in how they would solve them.

Analysis of quantitative results. Again, no gross quantitative differences between the two skill groups were produced. For example, there were no significant differences in the number of categories produced by each skill group (both groups averaged about 8.5 categories), and the four largest categories produced by each subject captured the majority (about 77%) of the problems. There was also little difference in the amount of time it took experts and novices to sort the problems, although experts tended to take slightly longer

time about 40 seconds per problem (discarding one outlier), whereas novices took about 37 seconds per problem.

The absence of gross quantitative differences in measures such as number of categories, number of largest categories, and time to categorize confirms the notion that there are no fundamental capacity differences between experts and novices. That is, the novices are not inherently slower, for example, nor do they have limited abilities to discriminate the problems into eight categories. The lack of a general quantitative difference points to the necessity of examining the qualitative differences.

Analysis of qualitative results. If we examine the nature of the categories into which experts and novices sorted the problems, they are qualitatively dissimilar. This difference can be most dramatically seen by observing the two pairs of problems that the majority of the subjects of each skill group sorted together. Figure 4 shows two pairs of problems that eight out of eight novices grouped together as similar. These problems have noticably similar "surface structures." By "surface structures," we mean either (a) the objects referred to in the problem (such as a spring or an inclined plane), (b) the keywords that have physics meaning (such as center of mass or friction), or (c) the physical configuration that involves the interaction of several object components (such as a block on an inclined plane).

The suggestion that these surface structures are the bases of the novices' categorization can be further confirmed by examining subjects' verbal justifications for the categories, which are

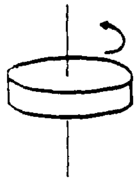
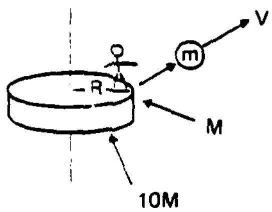
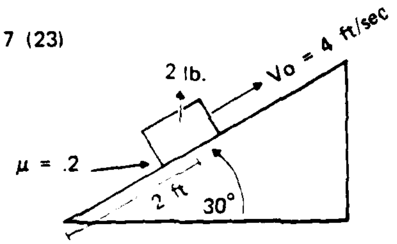
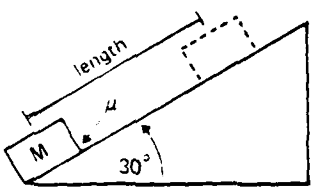
Diagrams Depicted from Problems Categorized by Novices within the Same Groups	*Novices' Explanations for Their Similarity Groupings*
Problem 10 (11) 	*Novice 2:* "*Angular* velocity, *momentum*, circular things" *Novice 3:* "*Rotational* kinematics, *angular* speeds, *angular* velocities" *Novice 6:* "Problems that have *something rotating; angular* speed"
Problem 11 (39) 	
Problem 7 (23) 	*Novice 1:* "These deal with blocks on an *incline plane*" *Novice 5:* "*Inclined plane* problems, coefficient of *friction*" *Novice 6:* "Blocks on *inclined planes* with angles"
Problem 7 (35) 	

Figure 4. Examples from novices' problem categories. Problem numbers represent chapter and problem number from Halliday and Resnick (1974)

presented on the right-hand column of Figure 4. The novices' explanations indicate that they grouped the top two problems together because they both involved "rotational things" and the bottom two together because they involved "blocks on an inclined plane."

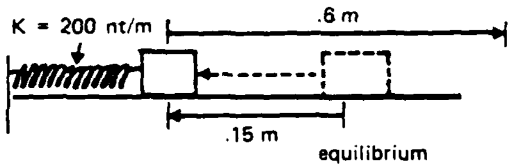
For experts, surface features do not seem to be the basis for categorization. There is neither a similarity in the keywords used in the problem statements, nor in the visual appearance of the diagrams for the problems, as shown in Figure 5. No similarity is apparent in the equations used for the problems that are grouped together by the majority of the experts. The similarity underlying the experts' categorization can only be detected by a physicist. It appears that the experts classify according to the major physics principles (or fundamental laws) governing the solution of each problem (sometimes referred to as the *solution method*). The top pair of problems in Figure 5 can be solved by the application of the Conservation of Energy Law while the bottom pair is better solved by the application of Newton's Second Law (F=MA). The verbal justifications of the subjects confirm this analysis. We might refer to the principles underlying a problem as the "deep structure" of the problem, which is the basis by which experts categorize problems.

In sum, the results of this study uncover several facets of problem solving that were not observable from protocol analyses. First, through a sorting task, it became apparent that categories of problems exist. These categories probably correspond to problem schemas, that is, unified knowledge that can be used to solve a particular type of problem. Second, category membership can be

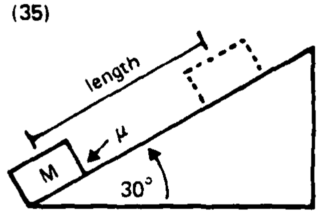
Diagrams Depicted from Problems Catergorized by Experts within the Same Groups

Experts' Explanations for Their Similarity Groupings

Problem 6 (21)



Problem 7 (35)

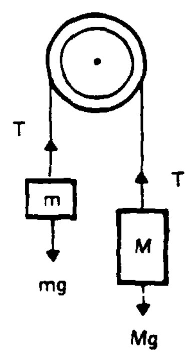


Expert 2: "*Conservation of Energy*"

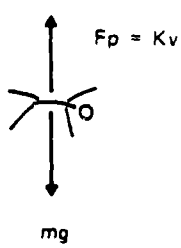
Expert 3: "*Work-Energy Theorem*. They are all straight-forward problems."

Expert 4: "These can be done from energy considerations. Either you should know the *Principle of Conservation of Energy*, or work is lost somewhere."

Problem 5 (39)



Problem 12 (23)



Expert 2: "These can be solved by *Newton's Second Law*"

Expert 3: "F = ma; *Newton's Second Law*"

Expert 4: "Largely use F = ma; *Newton's Second Law*"

Figure 5. Examples from experts' problem categories. Problem numbers represent chapter and problem number from Halliday and Resnick (1974).

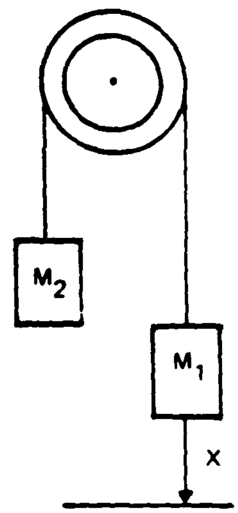
determined rather quickly, between 35-45 seconds. This is the amount of time we initially alloted to the qualitative analysis episodes of problem solving. Third, the results also imply that within 45 seconds, the experts at least, can already perceive the solution method applicable to the problem. The possibility that such categorization processes may occur during problem solving is never evident from the problem solving protocols, because there was never any cause for solvers to mention either the principle underlying a problem or the surface structure of the problem. Only through an alternative task, such as sorting, are we able to detect the presence of categories that may be related to solution methods.

Study Three: Sorting Specially Designed Problems

If the interpretation of the previous sorting results is accurate, then one should be able to replicate the findings, and further, to predict how a given subject of a specific skill level, might categorize a given problem. In this study, we specially designed a set of 20 problems to test the hypothesis that novices are more dependent on surface features whereas experts focus more on the underlying principles. Table 2 shows the problem numbers and the dimensions on which they were varied. The left column indicates the major objects that were used in the problem; the three right headings are the solution methods (or the basic laws) that can be used to solve them. Figure 6 shows an example of a pair of problems (corresponding to problems 11 and 18 in Table 2), which contain the same surface structure but different deep structure. In fact, the problems are

No. 11 (Force Problem)

A man of mass M_1 lowers himself to the ground from a height X by holding onto a rope passed over a massless frictionless pulley and attached to another block of mass M_2. The mass of the man is greater than the mass of the block. What is the tension on the rope?



No. 18 (Energy Problem)

A man of mass M_1 lowers himself to the ground from a height X by holding onto a rope passed over a massless frictionless pulley and attached to another block of mass M_2. The mass of the man is greater than the mass of the block. With what speed does the man hit the ground?

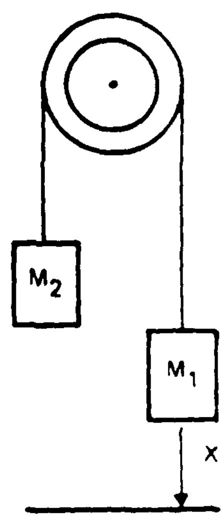


Figure 6. Sample problems.

identical except for the question asked. From the results of Study Two, we predicted that the novices would group together problems with similar surface features, such as the two problems shown in Figure 5, whereas experts would not. Experts, instead, would group together problems that have similar deep structure, regardless of the surface features. Intermediate subjects might exhibit some characteristics of each skill group.

The results confirmed our previous interpretations. Our one novice, who had completed one course in mechanics, grouped strictly on the surface structures of the problems. Table 3 shows his problem categories, and the explanations he provided for his groups. First of all, if one scans only the verbal justification column (far right), it is evident that, except for the fourth group, where he mentioned "Conservation of Energy," a physics principle, the remaining categories were all described by either physics keywords (such as "velocity problems"), or the actual physical components contained in the problem ("spring"). And indeed, he collapsed problems across the physics laws. For Group 5 (Table 3), problem 18 is obviously solvable by the Force Law, whereas problem 7 is solved by the Energy Law (see Table 2 again). The only category for which he made any reference to a physics principle is Group 4, which he described as a "Conservation for Energy" category. However, this is to be distinguished from the expert's labeling of "Conservation of Energy." This novice only labels those problems as "Conservation of Energy" when the term "Energy" is actually mentioned in the problem statements themselves, as was the case here.

In contrast, the expert's classifications are all explained by the underlying principles, such as Conservation of Angular Momentum, Conservation of Energy, etc. (see Table 4). Furthermore, as predicted, the expert collapsed problems across the surface similarities. For example, for Group 3, problem 1 is basically a spring problem, and problem 4 is a collision problem.

Table 5 shows the groupings of an advanced novice (an intermediate). His categorizations of the problems are characterized by the underlying physics principle in an interesting way. These principles are qualified and constrained by the surface components present in the problems. For example, instead of classifying all the Force problems together (Groups 4, 6, and 7), as would an expert, he explicitly separated them according to the surface features of the problems. That is, to him, there are different varieties of Force problems, some containing pulleys, some containing springs, and some containing inclined planes.

To summarize this study, we were able to replicate the initial finding that experts categorize problems by physics laws, whereas novices categorize problems by the literal components. If we assume that such categories reflect knowledge schemas, then our results from the person at the intermediate skill level suggest that with learning, there is a gradual shift in organization of knowledge, from one centering on the physical components, to one where there is a combined reliance on the physical components and physics laws, to finally, one primarily unrelated to the physical components.

Study Four: Hierarchical Sorting

The results of the previous two sorting studies strongly suggest that the problem categories of the experts are different from those of novices. That is, we assume that the differences lie not only in the "category labels" that subjects of different skills prefer to use. We assume that problem categories corresond to problem schemas and, theoretically, schemas can have embedded in them subschemas, and be embedded in higher-level or super-schemas. Hence, if we can identify some similarity of the contents of schemas at different levels for individuals of different skills, then perhaps we will have converging evidence that the schemas of the novices and experts are indeed different, and that their schemas might be the same when different levels are compared.

To test this assumption, a hierarchical sorting task was designed by Christopher Roth. In this task, subjects were first asked to sort the problems in the same manner as in the previous two studies. Then, their groups, which they had initially sorted, were returned to them, and they were asked to further subdivide each group, if they wished. The sorting of each group was conducted in a depth-first manner. When all the discriminations of each group were completed, they were also asked to combine their initial groups, until they no longer wished to make any further combinations. Subjects' rationale for each group that they made was also recorded.

Sixteen subjects were run. They ranged from graduate students (experts) to fourth year physics and chemical engineering majors (intermediates) to A and C students (novices) who had taken courses in

physics (mechanics and electricity and magnetism). A sample of these subjects' data will be discussed.

The 40 problems used in this study were selected from Halliday and Resnick, covering the chapters 5-12 of the text (as in Study Two), which is the minimumm amount of material typically covered in a first year mechanics course.

There are two aspects of the data to examine: the contents of the groups, and the tree structures. We believe that the most naive structures are those generated by the novice C students (R.R. and J.T.), as shown in Figure 7, top two panels. The circular nodes represent the groups from the initial sort, and the numbers inside the nodes indicate how many problems are in that group. The square nodes beneath the circular nodes are the groups formed when the problems were further discriminated, and the triangular nodes above the circular nodes indicate the combinations. The tree structures of these two novices have three distinct characteristics that none of the other more skilled subjects exhibited. First, the initial groups (circular nodes) have a greater than average number of categories. (Eight categories is the averge number derived from Study Two.) The second characteristic is that they either cannot make further discriminations (Novice R.R.), suggesting that their categories are already at the lowest level, or they make such fine discriminations (Novice J.T.) that each problem is in a category by itself. This is reminiscent of the chess results. Beginner chess players have chunks consisting of one or two pieces. The nature of the initial categories is physical configurations, much like what was found in Study Two,

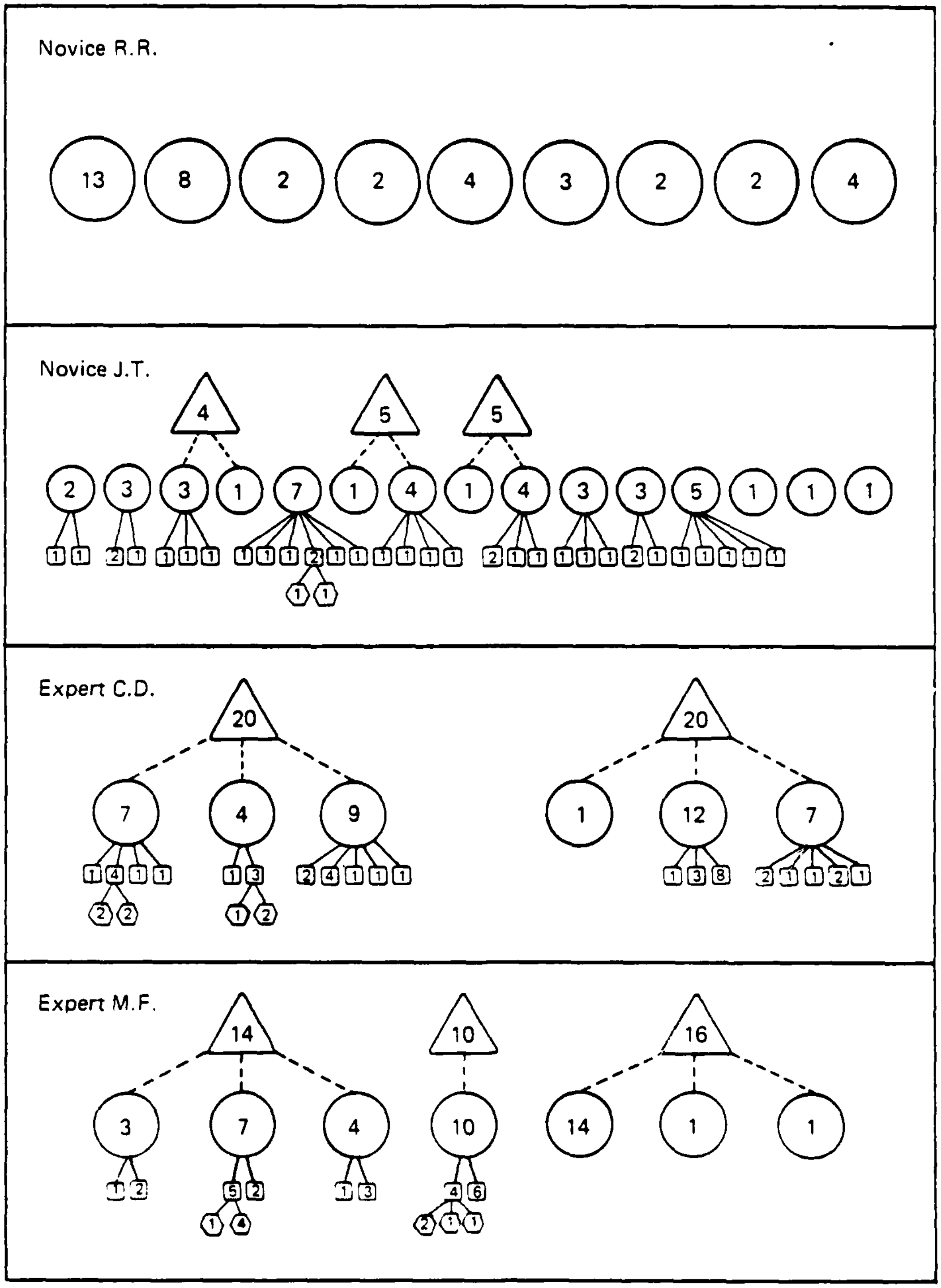


Figure 7. Groupings made by novices and experts on a hierarchical sorting task. Circular nodes are the preliminary groups made, squares and hexagons are subsequent discriminations, and triangles are the combinations.

such as "gravity," "pulley with weight," etc. When the novice (J.T.) breaks the categories down so that each problem is a category, the descriptions of these categories are very specific, and still bound to the physical configuration. For example, one of the initial categories of Novice J.T. is "tension in rope." When that category was further broken down, one of them was specified as "tension with two blocks on incline" and another was "tension with two blocks and pulley on incline."

The most sophisticated tree structures of the experts are shown in the lower two panels of Figure 7. The initial circular nodes are generally the different varieties of physics principles, much like those uncovered in Study Two. For Expert C.D., one group of circular nodes contains Conservation of Energy, Conservation of Momentum, and Conservation of Angular Momentum, and the other group of three are F=MA, F=MA to find the Resultant Force, and Simple Harmonic Motion. Each group of three (circled) categories was further collapsed to two superordinate categories: Conservation Laws and Equations of Motion. The subordinate categories for the same subject are generally discriminations based on physical configurations, such as "tension problems." Hence, from our limited analyses, we could hypothesize that the subordinate categories of the experts correspond to the initial categories of the novices. Although this study is not definitive in hypothesizing that experts' categories are at a higher level than novices' categories, additional data from Study Five will converge on the same notion.

The results of this study can also be interpreted in the framework proposed by Rosch (1978) of "basic" categories. The term "basic" can be used loosely to mean the preferred or dominant categories to which problems were divided by the subjects. Hence, one could say that the "basic" categories of the novices correspond to the subordinate categories of the experts.

Studies of the Knowledge Base

If the knowledge bases of the experts are different from those of the novices, in what ways are they organized differently, and in what way does the knowledge of experts and novices enhance and hinder their problem solving processes? These questions, coupled with the results of the categorization studies, lead us to an examination of the knowledge bases. The categorization studies show that without actually solving the problems, and in less than 45 seconds, experts were able to encode the problem into a deep level of representation, one that enables them to grossly determine the solution method applicable to the problem. We speculate that such encoding skill necessarily reflects the knowledge base differences between experts and novices. The next set of studies asks to what extent and in what ways are the knowledge bases of the novices less complete and coherent than the experts.

Study Five: Summaries

With these questions in mind, we attempted to capture what subjects knew about physics, independent of a problem solving content.

One simple approach was to ask subjects to summarize a chapter. This should reveal the knowledge they would have on a particular topic. We selected chapter five on particle dynamics of *Halliday and Resnick* (1974), because it was the knowledge in this chapter that subjects in the first protocol study needed in order to solve those five problems correctly. Furthermore, this chapter introduced Newton's three laws, which could be a common theme of the chapter that all subjects might mention during their summaries, so that we would be able to do some comparisons.

We asked four experts (two college professors, one postdoc who had never taught lower division physics, and one fifth-year graduate student who had often taught lower division physics) and four undergraduates (who had just completed the introductory physics course with a B grade, using Halliday and Rensick as a text) to review the chapter for five minutes, then summarize out loud the important concepts of the chapter. Subjects were run individually. Fifteen minutes were allotted for the summary. The book was also available to them while they summarized, so that any limitation in their summaries could not be attibuted to a retrieval problem. (Then they were all asked to solve a single problem taken from Chapter 5. These problem solving protocols provided the data for discussing the frequency of diagram drawing mentioned in Study One.)

We began again by looking at various quantitative measures, such as the length of the summaries, the number of quantitative relations mentioned in the summaries, and so on. Cursory examination of the data again suggested that there were no skill differences in any of

these quantitative measures. We then turned to an examination of the content of the summaries. Since every subject mentioned Newton's three laws of motion, we compared what they said about two of them.

The top of Table 6 states Newton's Third Law, and the bottom of the table shows one possible way of breaking down the law into its subcomponents. Using these subcomponents as a scoring criterion, we could analyze the summaries of the experts and novices, and see what proportion of the subcomponents were mentioned by each skill group. Such results are shown in Table 7. The X's in the table show the subcomponents of the law that were mentioned by each subject. On the bottom of this table are samples of protocols of a novice and an expert. It is clear from Table 7 that experts in general make more complete statements about the physical laws than novices, even though the textbook was available for them to use. Table 8 is another instance of a similar analysis of Newton's First Law. Again, experts mentioned on the averge three subcomponents, whereas novices tended to mention on the average at most two subcomponents. It is also interesting to note that Expert S.D.'s performance in Table 8 is most "novice-like," perhaps because he did not have any experience teaching mechanics.

The summaries of experts and novices on a given chapter from a physics text indicate that experts do have more complete information on physics laws than novices. This is not surprising in the sense that one would expect experts to know more. On the other hand, it is surprising because the students have been taught this knowledge and had the book available to them. One would hope that, after

instruction, the students have mastered at least the declarative knowledge of the laws of physics, however, one obvious deficiency of novices is that they had not. One cannot automatically assume that all students have mastered the prerequisite knowledge needed for solving problems. Nor can we assume that the novices' deficiencies lie mainly in the inadequate strategies or procedural knowledge that improves with experience in solving problems.

Up to this point, our data show that novices are deficient in three aspects of knowledge. First, very good students, as Study One shows, make errors in problem solving only when they have either generated the incorrect inferences or failed to generate the correct inference during the initial encoding or representation-generation stage of problem solving. We attribute the generating of the wrong inference to incomplete knowledge in the data base, so that the appropriate inference (the right link between certain nodes in the semantic network; Greeno & Riley, 1981) could not be made. Second, we discovered that, whether novices and experts have the same knowledge base or not, it is organized differently. That is, we can view the knowledge of problem types as schemas, and the experts' schemas center around the physics principles, whereas the novices' schemas center around the objects. Finally, a third deficiency in the novices' knowledge base, at least for B students, is that they lack certain fundamental knowledge of physics principles.

These three deficiencies in the knowledge base that we have already identified are general in the sense that we do not have a good grasp of exactly what knowledge is missing from the novices' data base

(except for the summary study), nor do we have any means for comparing the knowledge bases. And most importantly, we have tapped only the declarative knowledge that the subjects have. The next study attempts to be more detailed in assessing the knowledge that subjects do have, provides a means of comparing the knowledge bases between subjects, and begins to look at the use of procedural knowledge, since it is the procedural knowledge that will ultimately determine how well a person can solve a problem.

Study Six: Elaboration Study

In this study, we were interested in the knowledge associated with certain physics concepts. These are concepts generated by the category descriptors provided by the subjects in the sorting studies. We view these concepts as labels designating schemas. Hence, the purpose of the present study was to uncover what knowledge is contained in the schemas of experts and novices. From the sorting studies, we concluded that the schemas of the experts are principle-oriented, whereas the schemas of the novices are object-oriented. What we needed to know now is how the schemas of the two skill groups differ. Do the schemas of the experts contain more information, a different kind of information? Are the schemas of the novices subschemas of the expert schemas? This study addressed these issues.

Two experts (M.G., M.S.), both graduate students, and two novices (H.P., P.D.) were asked to elaborate on a selected sample of 20 prototypical concepts that subjects in the sorting studies had used to

describe their classifications. Figure 8 gives a frequency count of those category labels that were used by the experts and novices in Study Two. The sample of 20 used in this study ranged from those provided by experts (e.g., Force Law), to those provided strictly by novices (e.g., inclined plane). Subjects were presented with each concept individually, and given three minutes to tell everything they could think of about it, and how a problem involving the concept might be solved.

We use two ways to analyze the contents of these elaboration protocols. One way is to depict the contents of the protocol in terms of a node-link network, where the nodes are simply key terms that are mentioned that are obvious physics concepts. The links are simply unlabeled relations that join the concepts mentioned contiguously. Using this method, the networks of a novice's (H.P.) and an expert's (M.G.) elaboration of the concept "inclined plane" are shown in Figures 9 and 10. Since we view each of these concepts as representing a potential schema, the related physics concepts mentioned in the inclined plane protocol can be thought of as the variables (slots) of the schema. For example, in Novice H.P.'s protocol, his inclined plane schema contains numerous variables that can be instantiated, including the <u>angle</u> at which the plane is inclined with respect to the horizontal, whether there is a <u>block</u> resting on the plane, and what are the <u>mass</u> and <u>height</u> of the block. Other variables mentioned by the novice include the <u>surface property</u> of the plane, whether or not it has <u>friction</u>, and if it does, what are the <u>coefficients of static</u> and <u>kinetic friction</u>. The novice also

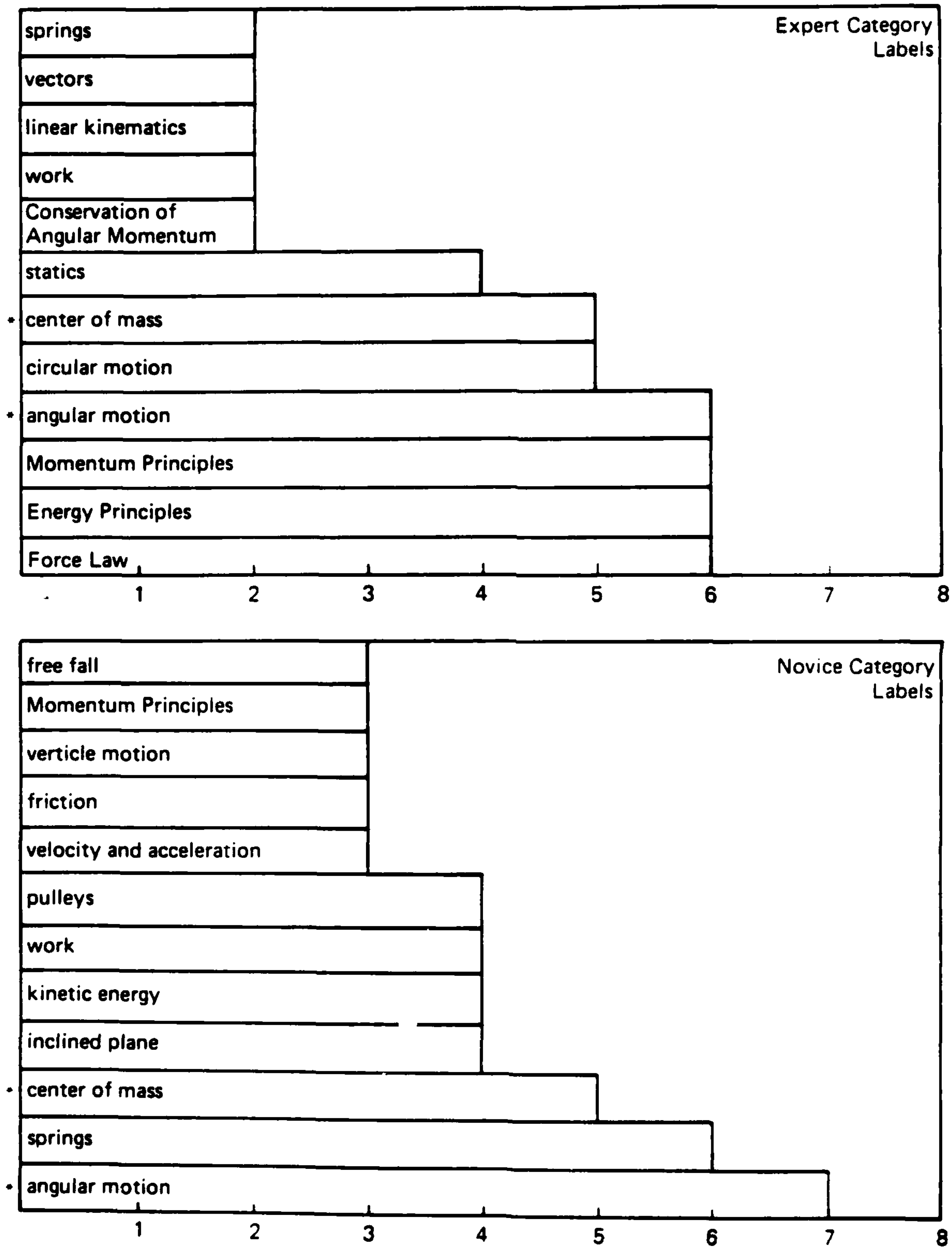


Figure 8. Frequency of use of category labels by eight experts and eight novices. Asterisks indicate labels used by both groups of subjects.

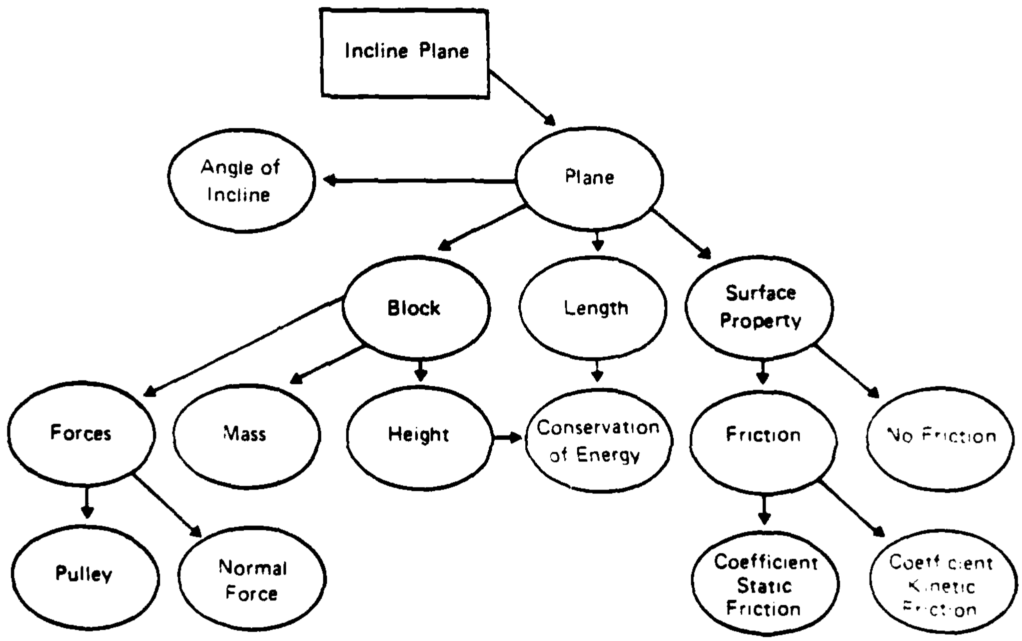


Figure 9. Network representation of Novice H.P.'s schema of an inclined plane.

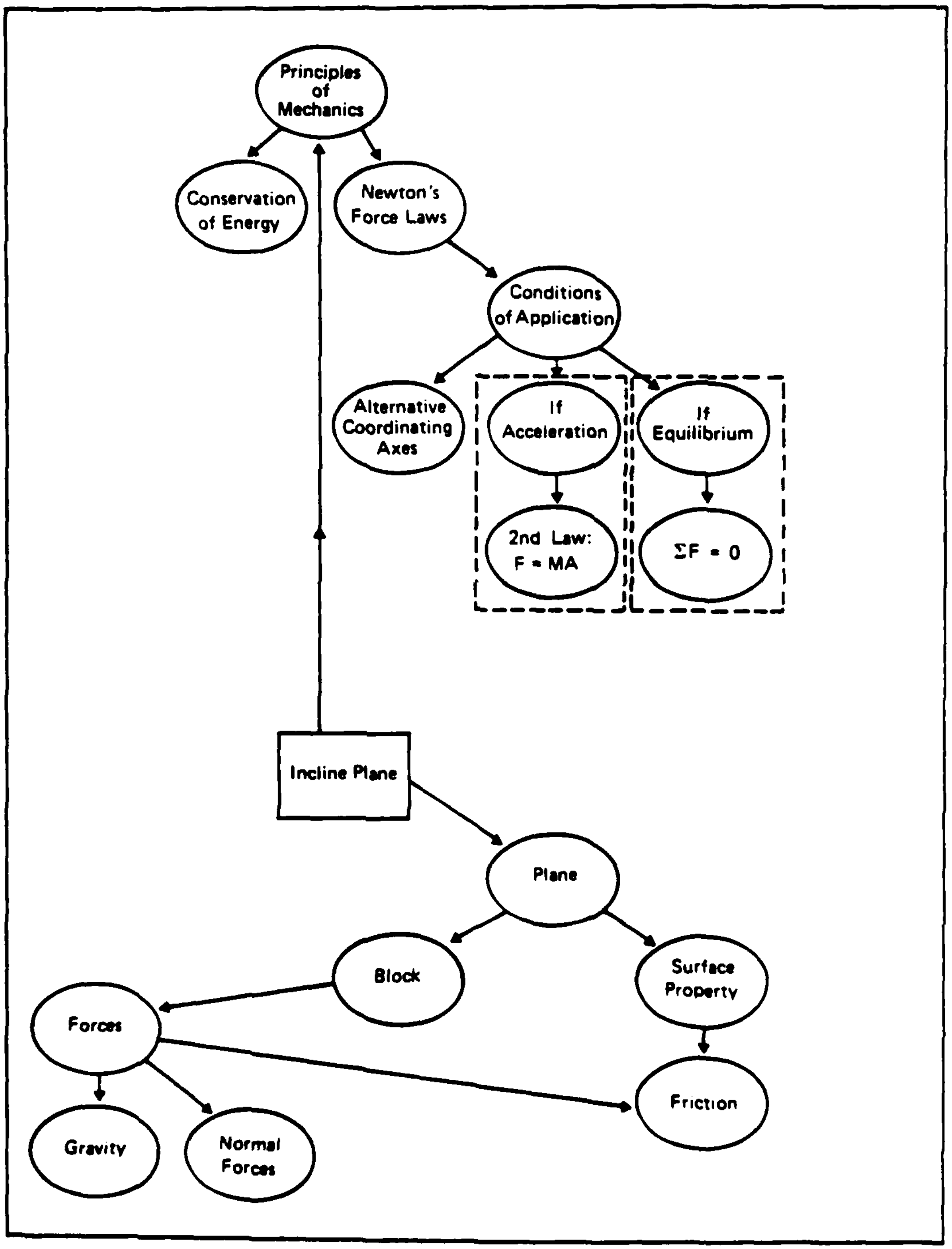


Figure 10. Network representation of Expert M.G.'s schema of an inclined plane.

discussed possible forces that may act on the blcok, such as possibly having a pulley attached to it. He also discussed, at the end, the pertinence of Conservation of Energy. However, his mentioning of the Conservation of Energy principle was not elicited as an explicit solution procedure that is applicable to a configuration involving an inclined plane, as is the case with the expert, as will be seen later. Hence, in general, one could say that the "inclined plane" schema that the novice possesses is quite rich. He knows precisely what variables need to be specified, and he also has default values for some of them. For example, if friction was not mentioned, he probably knows that he should ignore friction. Hence, with a simple specification that the problem is one involving an inclined plane, he can deduce fairly accurately what are the key components and entities (such as friction) that such a problem would entail.

The casual reference to the underlying physics principle, Conservation of Energy, given by the novice in the previous example, contrasts markedly with the expert's protocol, in which she immediately makes an explicit call to two priciples which take the status of procedures, the Conservation of Energy Principle, and the Force Law (Figure 10). (In Riley & Greeno's 1981 terminology, they would be considered calls to action schemata.) We characterize them as procedures (thus differentiating them from the way the novice mentioned a principle) because the expert, after mentioning the Force Law, continues to elaborate on the condition of applicability of the procedure, and then provides explicit formulas for two of the conditions (enclosed in dashed rectangles in Figure 10). (She also

explained the conditions of applicability of Conservation of Energy, but did so during other segments of the study.) After her elaboration of the principles and the conditions of applicability of one principle to inclined plane problems (depicted in the top half of Figure 10), Expert M.G. continued her protocol with descriptions of the structural or surface features of inclined plane problems, much like the descriptions provided by Novice H.P. (see Figure 9). Hence, it seems that the knowledge common to subjects of both skill groups pertains to the physical configuration and its properties, but that the expert has additional knowledge relevant to the solution procedures based on major physics laws.

Another perspective on the difference between the novice's and expert's elaborations of "inclined plane" is to look at the description that Rumelhart (1981) ascribes to schemas of inactive objects. That is, an "inclined plane" is viewed by the novice as an inactive object, so that it specifies not actions or event sequences, but rather, spatial and functional relationships characteristic of "inclined planes." Because novices may view inclined plane as an object, they thus cite the potential configuration and its properties. Experts, on the other hand, may view an inclined plane in the context of the potential solution procedures; that is, not as an object, but more as an entity that may serve a particular function.

An alternative way to analyze the same set of protocols is to convert them directly into "production rules," or IF-THEN rules (Newell, 1973). To do so, a simple set of conversion rules can be used, such as when the protocols manifest an IF-THEN or IF-WHEN or

WHEN-THEN structure. This transformation is quite straightforward, and covers a majority of the protocol data. Tables 9 and 10 depict the same set of protocols that were previously analyzed in the form of node-link structures. What is obvious from such an analysis is that the experts' production rules contain explicit solution procedures, such as "use F=MA," or "sum all the forces to 0." None of the novices' rules depicted in Table 10 contain any actions that are explicit solution procedures. Their actions can be characterized as attempts to find specific unknowns, such as "find mass" (see rules with asterisks in Table 10).

We alluded to an important difference between the way Conservation of Energy was mentioned by novice H.P. versus expert M.G. The present analysis makes this difference more transparent. The difference lies in the *observation* that the novice's statement of Conservation of Energy (Rule 8 in Table 10) was part of a description of the condition side of a production rule, whereas the statement of this principle by both experts (Table 9 see asterisks) is described on the action side of the production rules.

In Figure 10 on the elaboration of an inclined plane, we stressed the observation that the Expert mentioned the conditions of applicability of the Force Law (the statements in the dashed enclosures). This points to the presence of not only explicit procedures in the experts' repertoires, but also of explicit conditions for when a specific procedure applies. Another analysis supports this difference. We examined all statements made by the two experts and the two novices throughout the protocols of the entire set

of 20 concepts, and recorded all statements made about Conservation of Energy. Nearly half of each expert's statements (10 out of a total of 22 for Expert M.S., 9 out of a total of 21 for expert M.G.) were specifying the conditions under which Conservation of Energy could be used. For example, the following are two quotes, one from each subject.

Expert M.S. - "If the (inclined) plane is smooth, of course then you could use Conservation of Mechanical Energy to solve the problem. If it's not smooth, then you've got to take into account the work done by frictional forces."

Expert M.G. - "Energy conservation can also be used (in a collision problem) but only for an elastic collision because no heat is produced."

The novices on the other hand, made only one such statement between them (1/22 for H.P., 0/13 for P.D.).

In sum, this study shows that the contents of the schemas are different for the novices and the experts. First, for an object schema, both experts and novices possess the fundamental knowledge about the configuration and their properties; but the experts possess additional knowledge, which may be viewed also as activating higher level schemas (Rumelhart, 1981) that are relevant to the principle. Second, the schemas of the experts contain more procedural knowledge. That is, they have explicit procedures, which may be thought of as the action side of the productions. Finally, the experts' schemas contain much more knowledge about the explicit conditions of applicability of the major principles underlying a problem. Hence, this study, coupled

with the Summary Study, emphasizes the impoverished nature of novices' schemas, which can seriously hinder their problem solving success.

Studies to Identify the Key Features of Problems

The previous studies have suggested that novices in general, have deficient knowledge in a variety of ways (perhaps with the exceptions of A students). It is also important to ascertain whether the difficulties novices encounter in problem solving lie also in their inability to identify the relevant cues in the problem as is the case with poor chess players. The common finding in chess research is that the poorer players have greater difficulties seeing the meaningful patterns on the chess board. The ability to perceive the relevant chess-board patterns reflects the organization of the chess knowledge in memory. Hence, we need to determine whether novice and expert problem solvers both have the ability to identify the relevant cues in a problem, and if so, how this ability affects problem solving. From the studies we have already discussed, we speculate that the difficulties novices have derive from their inability to generate the appropriate knowledge from the relevant cues.

Study Seven: Basic Approach

In this study, designed and carried out by Paul Feltovich, we were interested in knowing about the features that help a subject decide on a "solution method," which can be interpreted as one of the three major principles (Conservation of Energy, Conservation of Momentum, and Force Law) that can underlie a mechanics problem of the

kind we use. Putting it another way, we are attempting to determine the problem features that subjects could have used in the eliciting of their category schemas, if the "solution methods," at least for the experts, may be viewed as their schemas of problem types (see Study Three).

Subjects in this study were asked to do three things. First, they were to read the problem statement, and think out loud about the "basic approach" that they would take to solve the problem. "Basic approach" was not further defined for them. Second, they were asked to re-state the "basic approach" explicitly in one concise phrase. Finally, they were asked to state the problem features that led them to their choice. We will focus predominantly on the last aspect of this study. Additional details can be gathered from Chi, Feltovich, and Glaser (in press). The subjects were two physicists (J.L., V.V.) who had frequently taught introductory mechanics, and two novices (P.D., J.W.) who had completed a basic college course in mechanics with an A grade. The problems used were the same 20 (described in Table 2) used for the sorting replication (Study Three).

Table 11 summarizes the key features cited by the experts and novices as contributing to their decisions about the "basic approach" to the solution of the problems. The numbers in the table show the frequency with which each feature was cited. A feature was included, for each skill group, only if it was mentioned at least twice (across the 20 problems), once by each subject or twice by one subject.

Analysis of these features shows, first of all, that there is essentially no overlap in the features mentioned by novices and

experts except for the object "spring." Second, the kinds of features mentioned as relevant by the novices are different from those identified by the experts. Novices, again, mention literal objects and key terms that are explicitly stated in the problem, such as "friction" and "gravity." This is consistent with the results of the categorization studies. Experts, on the other hand, identify features that can be characterized as descriptions of states and conditions of the physical situation, as described implicitly by the problem. In some instances, these are transformed or derived features, such as a "before and after situation" or "no external forces." Because these features are not explicitly stated in the problem, we refer to these as second-order features (or as we previously mentioned, generated tacit knowledge).

In sum, the most interesting finding of this study is that the features mentioned as relevant for suggesting a solution method are different for the experts and novices. Because the subjects used their own words to describe what the features are, there is often a lack of consensus concerning relevant features, particularly between the experts. In Table 11 for example, in 14 out of the 24 features cited, the experts did not refer to the same features, whereas this occurred only once for the novices (see the asterisks). This is consistent with the interpretation that novices must have geater consensus because they refer to the explicit key terms in the problem statement themselves. Experts, on the other hand, must necessarily show a great deal of individual differences because they transform the literal surface features into some second-order features, based on

their individual knowledge bases. However, even with such wide individual differences, there was a distinct characteristic to the experts' cited features that distinguished them from the novices' cited features.

Study Eight: Judging Problem Difficulty

Even though the experts cited the abstracted features as the relevant cues in the previous study, it is still possible that the experts transformed the same basic set of key terms as those identified by the novices. A direct way to ascertain whether subjects of different skills consider the same set of words as important, is to ask them to point out the important words in the problem statements. In this study, we presented six novices (approximately B students) and six experts (graduate students) the same set of 20 problems used earlier; and asked them to judge (using a 1-5 rating) how difficult a problem was to solve after reading the problem statement. We then asked them to circle the key words or phrases that helped them make that judgment. Finally, they were asked how those particular key words helped them make that decision.

The most striking finding is the extensive overlap between the cues that experts and novices identified as important for deciding on the difficulty of a problem. If anything, experts identified fewer cues as important, compared with the novices. Table 12 presents one of the 20 problems, broken down into eight propositions. There were, on the average, seven propositions per problem. The enclosed words were chosen by three or more of the six novices, and the asterisks

represent those that three or more experts selected. For 19 out of the 20 problems, the experts and the novices circled the same sets of words or phrases in the problem statements, which are embedded in 2.7 propositions, on the average. Only in seven of the 20 problems did the experts identify additional cues (about 1.6) whereas in 13 of the 20 problems, the novices identified an additional (2.1) cues as important. This result suggests, at least, that novices' difficulties in problem solving do not stem from their failure to identify the relevant cues.

The subjects' responses to both the questions of why these particular cues are important and how they help them make decisions, were classified according to the following categories: (1) whether the cues refer to one of the three fundamental principles ("the cues tell me to use Energy Conservation"), (2) whether the cues refer to some surface feature of the problem, much like what novices refer to when they categorize problems, (e.g., Figure 8), (3) whether the cues bring their attention to some characteristic of the problem that is not related to physics ("it is difficult to visualize," or "it has many concepts"), or (4) whether the cues elicit some reasons that are unrelated to the specific problem (the problem is difficult "because I have never solved it before," or "it has a lot of words"). Table 13 is a breakdown of experts' and novices' reasons for why a problem was ·dged difficult or easy, along with samples of quotes. Consistent with our previous findings, experts, much more often than novices, rely on the underlying physics principle when judging the difficulty of a problem (e.g., "compressing spring tells me to think Energy".

AD-A100 138 PITTSBURGH UNIV PA LEARNING RESEARCH AND DEVELOPMEN--ETC F/G 5/10
EXPERTISE IN PROBLEM SOLVING.(U)
MAY 81 M T CHI, R GLASER, E REES N00014-78-C-0375
UNCLASSIFIED TR-5 NL

2 OF 2
AD A
00138

END
DATE
FILMED
7-81
DTIC

They both rely equally often o[illegible]lem characteristics, such as whether a problem involves friction or the center of mass. However, novices are much more likely than experts to rely on superficial nonphysics aspects of a problem to make their judgments (the third category in Table 13) such as whether "it is abstractly phrased," and "it has a lot of words." Finally, the novices often introduce reasons for why a problem is difficult that are not specific to a given problem, such as "I have never done problems like this before."

When inferences were generated in the protocols of problem solving (Study One), and when second-order features were identified (Study Seven), we speculated that such tacit knowledge was generated from the literal key terms in the problem statement. Now, we can verify some of these speculations directly, by examining some of the reasons that subjects gave for how some particular key terms that they circled contributed to their judgment of problem difficulty. Table 14 presents examples of the kind of statements produced by experts. These statements of reasons can be judged to be inferences generated either directly from the literal terms in the problem, such as "frictionless, use Conservation of Momentum," or the inferences may be generated from a derived cue, such as "no dissipative forces." These correspond to the second-order features mentioned in the previous study.

Recall that the purpose of this task was to ask the experts and novices to judge problem difficulty. The experts, in general, were more accurate at judging the difficulty of a problem than novices. Accuracy was determined by comparing the ratings of problem

difficulties that subjects gave with our own assessment of how difficult a problem actually is to solve. The aforementioned examination of the reasons subjects gave for why a particular problem is difficult, and why those particular keywords were helpful in identifying a problem's difficulty (Table 13), suggest that novices are less accurate at judging a problem's difficulty because they rely heavily on nonphysics related or nonproblem related features to determine its difficulty. Obviously these are not the reliable factors to consider when one attempts to solve a problem.

In sum, even though the task of this study--requesting sources of problem difficulty--is slightly different from either a problem solving task, or tasks used in the other studies, such as sorting, we suspect that the features identified as relevant in this task are the same as those used in other tasks. Basically, the results show that the relevant and important key terms in a physics problem can be identified by novices quite accurately. In this sense, a physics problem is not analogous to a "perceptual" chessboard, in which case the beginner cannot pick out the relevant or important patterns. However, the similarity between a chess expert and a physics expert remains, and can be seen in their ability (compared to novices) to abstract the relevant tacit knowledge cued by the external stimuli. The chess masters' superior ability derives from the ability to abstract or impose a cognitive structure onto the pattern of black and white chess pieces. That is, novice chess players are just as capable as experts at perceiving the chess pieces per se. However, to "see" the relations among the pieces require the fitting of one's schemas

(perhaps) to the configuration of chess pieces. Similarly, the novice physicist is just as capable as the expert physicist at identifying the key terms in a problem statement. The difficulty resides in the novice's limited ability to generate inferences and relations not explicitly stated in the problem.

GENERAL DISCUSSION

The goal of this chapter has been to contribute to our understanding of high-level competence in complex domains of human knowledge. Expert individuals in various areas of knowledge perform remarkable intellectual activities, and cognitive psychlogists are on the threshold of understanding these feats of memory retrieval, rapid perception, and complex problem solving. Since intelligence is generally measured through tests that assess skill in acquiring new knowledge in scholastic settings, understanding the nature of the competence attained should shed light on this ability to learn.

Early in this chapter, evidence was provided for the necessity to focus on the organization and structure of knowledge, in both psychological and AI research. This trend toward understanding the influence of knowledge is relatively recent, in contrast to the earlier emphasis on search algorithms and other heuristics for deducing and retrieving information. The techniques and theories that evolved, such as means-end analysis, were intended to be independent of the particular data base, and as such, have proven to be valuable search heuristics that are generalizable across different tasks and knowledge domains.

The turn to a focus on the knowledge base was necessitated in part by the inability of psychological theories to model human capabilities solely on the basis of search heuristics, and in part by the limitations discovered in attempting to construct AI programs that would outperform humans, even though the computer's search capabilities are essentially limitless. Hence, the constraints of powerful search techniques, when they did not engage an organized knowledge structure, soon compelled researchers to develop theories and programs that took account of the role of knowledge structure.

The emphasis on the knowledge base has also changed the direction of research. Since knowledge has different degrees of structure, depending on an individual's experience, it was intuitively apparent that an important problem was how a particular knowledge base is structured. The obvious choice was to model the expert's knowledge, as was done most dramatically in a number of AI programs. This choice has also led to psychological investigations of developing structure of novices' knowledge, in contrast to the richly organized structure of experts' knowledge.

The research on problem solving generated by this new emphasis has revolved around understanding the processes of arriving at a solution, in the context of the knowledge available to a solver. In physics, this has led to the construction of numerous theoretical models that attempt to simulate the processes of problem solving, in particular, the knowledge that is necessary to generate a particular sequence of equations. Other theoretical models constructed by AI

researchers have put more emphasis on the representation of the problem in the context of the available knowledge.

The important issue of problem representation has also been recognized in the psychological research. It is conspicuous in protocols of problem solving in the form of "qualitative analysis" of the problem, which usually occurs early in the solution process. Most empirical findings to date have failed to explicate this initial "qualitative analysis" of the problem, although the consensus has been that a representation of the problem, constructed at this point, is a significant factor in driving the solution process. Numerous quantitative differences between the experts and novices have also been identified, such as solution speed, errors, and equation generation pattern. None of these measures, however, has succeeded in shedding much light on understanding the different problem-solving processes of experts and novices.

The research from our own laboratory has been oriented toward magnifying the representational "stage" of problem solving through techniques other than the analysis of problem-solving protocols. Our findings (Study One) have emphasized the point that solution protocols provide limited insights to the processes of representation, and further, produce quantitative measures that are difficult to interpret because they are subject to large individual differences. These individual differences are dictated by a variety of particular strategies that solvers adopt, such as generating a number of equations when one cannot think of a way to proceed. Through the use of a sorting task (Studies Two, Three, and Four), we were able to

uncover a potential source of representational difficulty for novices. If we assume that a problem is represented in the context of the available knowledge, then novices will undoubtedly have an incomplete and less coherent representation, because of the organization of their knowledge. Their knowledge is organized around dominant objects (such as an inclined plane), and physics concepts (such as friction) mentioned explicitly in the problem statement. Experts, on the other hand, organize their knowledge around fundamental principles of physics (such as Conservation of Energy) that derive from tacit knowledge not apparent in the problem statement. An individual's "understanding" of a problem has been explicitly defined as being dictated by knowledge of such principles (Greeno & Riley, 1981). Hence, during "qualitative analysis" of a problem, an expert would "understand" a problem better than a novice, because he "sees" the underlying principle.

A person's "understanding" of a principle can be evaluated in several ways (Greeno & Riley, 1981). One way is to have it stated explicitly, as was done by experts in the Summary Study (Study Five), and in the rationale they provided in the Sorting Studies (Two, Three, and Four). Another way is to analyze the nature of the categories into which individuals sort problems; this constitutes an implicit assessment of their "understanding" of principles. An alternative but consistent interpretation of the Sorting Studies is that experts and novices organize their knowledge in different ways. Experts possess schemas of principles that may subsume schemas of objects, whereas novices may possess only schemas of objects. Some support for this

conjecture was provided in both Study Four, on the hierarchical nature of the sorting categories, and in Study Six, on the elaboration of the contents of object and principle schemas. Once the correct schema is activated, knowledge (both procedural and declarative) contained in the schema is used to process the problem further. The declarative knowledge contained in the schema generates potential problem configurations and conditions of applicability for procedures, which are then tested against the information in the problem statement. The procedural knowledge in the schema generates potential solution methods that can be used on the problem. Experts' schemas contain a great deal of procedural knowledge, with explicit conditions for applicability. Novices' schemas may be characterized as containing sufficiently elaborate declarative knowledge about the physical configurations of a potential problem, but lacking in abstracted solution methods.

Our hypothesis is that the problem-solving difficulties of novices can be attributed mainly to inadequacies of their knowledge bases, and not to limititations in either the architecture of their cognitive systems or processing capabilities (such as the inability to use powerful search heuristics or the inability to detect important cues in the problem statement). This conjecture follows from several findings. First, similarity in the architecture of experts' and novices' cognitive systems is probably implied by the fact that there are generally no differences between experts and novices in the number of categories into which they prefer to sort problems, in the latency required to achieve a stable sort, and in a variety of other measures.

These quantitative measures point to the invariance in the cognitive architecture of experts and novices. Second, novices do show effective search heuristics when they solve problems using backward-working solutions. Thirdly, in our last set of studies (Studies Seven and Eight), we showed that novices are essentially just as competent as experts in identifying the key features in a problem statement. The limitation of the novices derives from their inability to infer further knowledge from the literal cues in the problem statement. In contrast, these inferences necessarily are generated in the context of the relevant knowledge structures that experts possess.

In concluding this chapter, we would like to speculate on the implications of the work and theory reported here for a conception of intelligence. The tests of intelligence in general use today measure the kind of intellectual performance most accurately called "general scholastic ability." Correlational evidence has shown that the abilities tested are predictive of success in school learning. Given this operational fact, these commonly used tests of intelligence are not tests of intelligence in some abstract way. Rather, if we base our conclusions on their predictive validity, we can conclude that they are primarily tests of abilities that are helpful for learning in present-day school situations. More generally, we can assume that these intelligence tests measure the ability to solve problems in school situations, which leads to learning. The problem-solving ability possessed by the expert learner is a result of experience with the domains of knowledge relevant to schooling.

If expertise in learning is the abiity for representing and solving school problems, then for a less intelligent learner, a problem representation may be in close correspondence with the literal details of a problem, while for a more intelligent learner, the representation contains, in addition, inferences and abstractions derived from knowledge structures acquired in past experiences. As a result of prior experience in various knowledge domains relevant to schooling, the representations required for solving school problems are more enriched, and contribute to the ease and efficiency with which learning problems are solved. We speculate further that the knowledge the expert learner brings to a problem would incorporate a good deal of procedural knowledge--how a knowledge structure can be manipulated, the conditions under which it is applicable, etc. Novice learners, on the other hand, would have sufficient factual and declarative knowledge about a learning problem, but would lack procedural skill and this would weaken their ability to learn from their available knowledge.

A knowledge-based conception of intelligence could have implications for how individuals might be taught to be more effective learners. Such an attempt would de-emphasize the possibility of influencing mental processing skill (i.e., developing better methods for searching memory). Improved ability to learn would be developed through a knowledge strategy in which individuals would be taught ways in which their available knowledge can be recognized and manipulated. Improvement in the skills of learning might take place through the exercise of procedural (problem-solving) knowledge in the context of

specific knowledge domains. To date, conceptions of intelligence have been highly process oriented, reminiscent of earlier notions of powers of mind. If, in contrast, one did take a knowledge-emphasis approach to the differences between high and low performers in school learning, then one might begin to conduct investigations of knowledge structure and problem representation in the way that we have begun to do in the expert-novice studies described in this chapter. This orientation might provide new insights into the nature of the expert performance we define as intelligence.

Acknowledgment

The research reported here was supported in part by Grant N00014-78-C-0375, NR 157-421 from the Office of Naval Research, and in part by the Learning Research and Development Center, University of Pittsburgh, which is funded in part by the National Institute of Education. Our research on physics problem solving represents a team effort. The person who took major responsibilities for any single study is recognized as the study is discussed. However, almost everyone worked to some extent on each study. We acknowledge particularly the efforts of our colleague Paul Feltovich. Our thanks also go to Andrew Judkis, Tom Laritz, and Christopher Roth. We are indebted to all the physics professors who generously contributed their time by participating in the study.

References

Akin, O. Models of architechtural knowledge. London: Pion, 1980.

Atwood, M. .E., Masson, M. E.JS., & Polson, P. G. Further explorations with a process model for water jug problems. Memory and Cognition, 1980, 8, 189-192.

Atwood, M. E., & Polson, P. G. A process model for water jug problems. Cognitive Psychology, 1976, 8, 191-216.

Belmont, J. M., & Butterfield, E. C. Learning strategies as determinants of memory deficiencies. Cognitive Psychology, 1971, 2, 411-421.

Bhaskar, R., & Simon, H. A. Problem solving in semantically rich domains: An example from engineering thermodynamics. Cognitive Science, 1977, 1, 193-215.

Brown, J. S., Collins, A., & Harris, G. Artificial intelligence and learning strategies. In H. O'Neil (Ed.), Learning strategies. New York: Academic Press, 1978.

Bundy, A. Will it reach the top? Prediction in the mechanics world. Artificial Intelligence, 1978, 10, 129-146.

Bundy, A., Byrd, L., Luger, G., Mellish, C., & Palmer, M. Solving mechanics problems using meta-level inference. Proceedings of the 6th International Joint Conference on Artificial Intelligence, Tokyo, 1979.

Byrd, L., & Borning, A. Extending Mecho to solve statics problems (Tech. Rep. DAI 137). University of Edinburgh, Department of Artificial Intelligence, 1980.

Champagne, A. B., Klopfer, L. E., & Anderson, J. H. Factors influencing the learning of classical mechanics. *American Journal of Physics*, 1980, *8*, 1074-1075.

Charness, N. Components of skill in bridge. *Canadian Journal of Psychology*, 1979, *33*, 1-50.

Chase, W. G., & Chi, M. T. H. Cognitive skill: Implications for spatial skill in large-scale environments. In J. Harvey (Ed.), *Cognition, social behavior, and the environment*. Potomac, Md.: Erlbaum, in press.

Chase, W. G., & Simon, H. A. Perception in chess. *Cognitive Psychology*, 1973, *4*, 55-81. (a)

Chase, W. G., & Simon, H. A. The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973. (b)

Chi, M. T. H. Short-term memory limitations in children: Capacity or processing deficits? *Memory and Cognition*, 1976, *4*, 559-572.

Chi, M. T. H. Knowledge structures and memory development. In R. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum, 1978.

Chi, M. T. H., Feltovich, P., & Glaser, R. Categorization and representation of physics problems by experts and novices. *Cognitive Science*, in press.

Chiesi, H., Spilich, G. J., & Voss, J. F. Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 1979, *18*, 257-273.

de Groot, A. Thought and choice in chess. The Hague: Mouton, 1965.

de Groot, A. Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), Problem solving. New York: Wiley, 1966.

de Kleer, J. Multiple representations of knowledge in a mechanics problem solver. Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977.

Egan, D. E., & Greeno, J. G. Theory of rule induction: Knowledge acquired in concept learning, serial pattern learning, and problem solving. In L. Gregg (Ed.), Knowledge and cognition. Hillsdale, N.J.: Erlbaum, 1974.

Egan, D., & Schwartz, B. Chunking in recall of symbolic drawings. Memory and Cognition, 1979, 7, 149-158.

Eisenstadt, M. & Kareev, Y. Aspects of human problem solving: The use of internal representation. In D. A. Norman & D. E. Rumelhart (Eds.), Explorations in cognition. San Francisco: Freeman, 1975.

Greenblatt, R. D., Eastlake, D. E., & Crocker, S. D. The Greenblatt chess program. Proceedings of the Fall Joint Computer Conference, 1967, 31, 801-810.

Greeno, J. G. Hobbits and orcs: Acquisition of a sequential concept. Cognitive Psychology, 1974, 6, 270-292.

Greeno, J. G. Process of understanding in problem solving. In N. J. Castellan, D. B. Pison, & G. R. Potts (Eds.), Cognitive theory (Vol. 2). Hillsdale, N.J.: Erlbaum, 1977.

Greeno, J. G. Natures of problem solving abilities. In W. K. Estes (Ed.), Handbook of learning and cognitive processes (Vol. 5). Hillsdale, N.J.: Erlbaum, 1978.

Greeno, J. G., & Riley, M. S. Processes and development of understanding. In F. E. Weinert & R. Kluwe (Eds.), Learning by thinking. West Germany: Kuhlhammer, 1981.

Halliday, D., & Resnick, R. Fundamentals of physics (2nd edition). New York: John Wiley & Sons, 1974.

Hayes, J. R., & Simon, H. A. The understanding process: Problem isomorphs. Cognitive Psycholgoy, 1976, 8, 165-190.

Hinsley, D. A., Hayes, J. R., & Simon, H. A. From words to equations: Meaning and representation in algebra word problems. In P. A. Carpenter & M. A. Just (Eds.), Cognitive processes in comprehension. Hillsdale, N.J.: Erlbaum, 1978.

Hunt, E., Lunneborg, C., & Lewis, J. What does it mean to be high verbal? Cognitive Psychology, 1975, 7, 194-227.

Jeffries, R., Turner, A. T., Polson, P. G., & Atwood, M. E. Processes involved in designing software. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, N.J.: Erlbaum, 1981.

Jensen, A. R. Reaction time and intelligence. In M. Friedman, J. P. Das, & N. O'Connor (Eds.), Intelligence and learning. New York: Plenum Press, in press.

Larkin, J. H. Problem solving in physics (Tech. Rep.). Berkeley: University of California, Group in Science and Mathematics Education, 1977. (a)

Larkin, J. H. Skilled problem solving in physics: A hierarchical planning model. Unpublished manuscript, University of California at Berkeley, 1977. (b)

Larkin, J. H. Processing information for effective problem solving. Engineering Education, 1979, 70(3), 285-288.

Larkin, J. H. Understanding, problem representations, and skill in physics. Paper presented at NIE-LRDC Conference on Thinking and Learning Skills, Pittsburgh, October 1980.

Larkin, J. H. Enriching formal knowledge: A model for learning to solve textbook physics problems. In J. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, N.J.: Erlbaum, 1981.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. Models of competence in solving physics problems. Cognitive Science, 1980, 4, 317-345.

McCloskey, M., Caramazza, A., & Green, B. Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. Science, 1980, 210(5), 1139-1141.

McDermott, J., & Larkin, J. H. Re-representing textbook physics problems. Proceedings of the 2nd National Conference, the Canadian Society for Computational Studies of Intelligence. Toronto: University of Toronto Press, 1978, 156-164.

McKeithen, K.B. Assessing knowledge structures in novice and expert programmers. Unpublished doctoral dissertation, University of Michigan, 1979.

Minsky, M., & Papert, S. Artifical intelligence. Condon Lectures, Oregon State System of Higher Education, Eugene, Oregon, 1974.

Neves, D. M., & Anderson, J. R. Knowledge compilation: Mechanisms for the automatization of cognitive skills. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, N.J.: Erlbaum, 1981.

Newell, A. Production systems: Models of control structures. In W. G. Chase (Ed.), Visual information processing. New York: Academic Press, 1973.

Newell, A., Shaw, J. C., & Simon, H. A. Chess-playing programs and the problem of complexity. IBM Journal of Research and Development, 1958, 2, 320-335.

Newell, A., Shaw, J. C., & Simon, H. A. Report on a general problem solving program for a computer. Information Processing: Proceedings of the International Conference on Information Processing. Paris: UNESCO, 1960.

Newell, A., & Simon, H. A. Human problem solving. Englewood Cliffs, N.J.: Prentice-Hall, 1972.

Novak, G. S., Jr. Representations of knowledge in a program for solving physics problems. Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977.

Paige, J. M., & Simon, H. A. Cognitive processes in solving algebra word problems. In B. Kleinmuntz (Ed.), Problem solving. New York: Wiley, 1966.

Perfetti, C. A., & Hogaboam, T. The relation between single word decoding and reading comprehension skill. Journal of Educational Psychology, 1975, 67(4), 461-469.

Polson, P. G., & Jeffries, R. Problem solving as search and understanding. In R. Sternberg (Ed.), Advances in the psychology of intelligence (Vol. 1). Hillsdale, N.J.: Erlbaum, 1981.

Pople, H. E. The formation of composite hypotheses in diagnostic problem solving: An exercise in synthetic reasoning. Proceedings of the 5th International Joint Conference on Artificial Intelligence, Cambridge, Mass., 1977.

Reed, S. K., & Simon, H. A. Modeling strategy shifts in a problem solving task. Cognitive Psychology, 1976, 8, 86-97.

Reitman, J. Skilled perception in Go: Deducing memory structures from inter-response times. Cognitive Psychology, 1976, 8, 336-356.

Riley, M. S., Greeno, J. G., & Heller, J. I. Development of word problem solving ability. In H. P. Ginsberg (Ed.), Development of mathematical thinking. New York: Academic Press, 1981.

Rosch, E. Principles of categorization. In E. Rosch & B. Lloyd (Eds.), Cognition and categorization. Hillsdale, N.J.: Erlbaum, 1978.

Rumelhart, D. E. Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Erlbaum, 1981.

Shannon, C. E. Programming a digital computer for playing chess. Philosophy Magazine, 1950, 41, 356-375.

Shavelson, R. J. Methods for examining reresentations of a subject-matter structure in a student's memory. Journal of Research in Science Teaching, 1974, 11(3), 231-249.

Shavelson, R. J., & Stanton, G. C. Construct validation: Methodology and application to three measures of cognitive structure. Journal of Educational Measurement, 1975, 12(2), 67-85.

Simon, D. P., & Simon, H. A. Individual differences in solving physics problems. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, N.J.: Erlbaum, 1978.

Simon, H. A. The functional equivalence of problem solving skills Cognitive Psychology, 1975, 7, 268-288.

Simon, H. A. The infomation storage system called "Human Memory." In M. R. Rosenzweig & E. L. Bennett (Eds.), Neural mechanisms of learning and memory. Cambridge, Mass.: MIT Press, 1976.

Simon, H. A., & Barenfield, M. Information processing analysis of perceptual processes in problem solving. Psychological Review, 1969, 76, 473-483.

Sloboda, J. A. Visual perception of musical notation: Registering pitch symbols in memory. Quarterly Journal of Experimental Psychology, 1976, 28, 1-16.

Thomas, J. C., Jr. An analysis of behavior in the hobbits-orcs problems. Cognitive Psychology, 1974, 6, 257-269.

Thro, M. P. Relationships between associative and content structure of physics concepts. Journal of Educational Psychology, 1978, 70(6), 971-978.

Voss, J. F., & Tyler, S. Problem solving in a social science domain. Unpublished manuscript, University of Pittsburgh, Learning Research and Development Center, 1981.

Waterman, D. A. Adaptive production systems. _Proceedings of the 4th International Joint Conference on Artificial Intelligence_, Georgia, USSR, 1975, 296-303.

Wilkins, D. Using patterns and plans in chess _Artificial Intelligence_, 1980, _14_, 165-203.

Wolf, T. A cognitive model of musical sight reading. _Journal of Psycholinguistic Research_, 1976, _5_, 143-171.

Wortman, P. M. Medical diagnosis: An information processing approach. _Computer and Biomedical Research_, 1972, _5_, 315-328.

Table 1

Solution Time (Sec), Number of Equations Generated, and Number of Diagrams Drawn, for Each Subject and Problem

Problems (No. of Subparts)		Problem 1 (1)	Problem 2 (2)	Problem 3 (2)	Problem 4 (3)	Problem 5 (2)	Mean[a]
Expert R.E.	Solution Time	225	625	555	590	585	516
	No. of Equations	6	8	12	9	14	9.8
	No. of Diagrams	3	4	4	1	2	2.8
Expert M.V.	Solution Time	(165)	(325) A,B	500	590	590	560
	No. of Equations	3	5	7	12	15	8.4
	No. of Diagrams	1	1	1	2	3	1.6
Novice C.H.	Solution Time	275	(585) A,B	(925) A	(835) A,B,C	(325) A,B[b]	275
	No. of Equations	7	10	12	19	8	11.2
	No. of Diagrams	3	3	5	3	3	3.4
Novice K.W.	Solution Time	200	105	(290) B	655	420	345
	No. of Equations	7	10	12	19	7	11.0
	No. of Diagrams	2	0	2	2	1	1.4

Note. Parentheses around the solution time indicate an incorrect solution. The following lette (s) indicates the incorrect part(s) of the problem.

[a] The mean solution time was calculated only for problems correctly solved.

[b] The subject attempted only Part A of this problem.

Table 2

Problem Categories

Surface Structure	Principles: Forces	Energy	Momentum (Linear or Angular)
Pulley with hanging blocks		20†	
	11	19†	
	14*	3*†	
Spring		7	
	18	16	1
			17+
		9	6+
Inclined Plane	14*	3*†	
		5	
Rotational	15		2
			13
Single hanging block	12		
Block on block	8		
Collisions (Bullet-"Block" or Block-Block)			4
			6+
			10+

Note. * Problems with more than one salient surface feature. Listed multiply by feature.
† Problems that could be solved using either of two principles, energy or force.
+ Two-step problems, momentum plus energy.

Table 3

Problem Categories and Explanations for Novice H. P.

Group 1:	2, 15	"Rotation"
Group 2:	11, 12, 16*, 19	"Always a block of some mass hanging down"
Group 3:	4, 10	"Velocity problems" (collisions)
Group 4:	13†, 17	"Conservation of Energy"
Group 5:	6, 7, 9, 18	"Spring"
Group 6:	3, 5, 14	"Inclined plane"
Groups 7, 8, 9 were singletons		

Note. * Problem discrepant with our prior surface analysis as indicated in Table 3
† Problems disrepant with our prior principles analysis as indicated in Table 3.

Table 4

Problem Categories and Explanations for Expert V. V.

Group 1:	2, 13	"Conservation of Angular Momentum"
Group 2:	18	"Newton's Third Law"
Group 3:	1, 4	"Conservation of Linear Momentum"
Group 4:	19, 5, 20, 16, 7	"Conservation of Energy"
Group 5:	12, 15, 9†, 11, 8, 3, 14	"Application of equations of motion" (F = MA
Group 6:	6, 10, 17	"Two-step problems: Conservation of Linear Momentum plus an energy calculation of some sort"

Note. † Problem discrepant with our prior principles analysis.

Table 5

Problem Categories and Explanations for Advanced Novice M. H.

Group 1:	*14*, *20*	"Pulley"
Group 2:	1, *4*, *6*, *10*, 12†	"Conservation of Momentum" (collision)
Group 3:	*9*, 13†, *17*, *18*†	"Conservation of Energy" (springs)
Group 4:	*19*, *11*	"Force problems which involve a massless pulley" (pulley)
Group 5:	*2*, *15*†	"Conservation of Angular Momentum" (rotation)
Group 6:	*7*†, *16*†	"Force problems that involve springs" (spring)
Group 7:	8, *5*†, *3*	"Force problems" (inclined plane)

Note. Italic numbers mean that these problems share a similar surface feature, which is indicated in the parentheses, if the feature is not explicitly stated by the subject.
† Problems discrepant with our prior principles analysis.

Table 6

Newton's Third Law and Its Decomposition

"To every action there is always opposed an equal reaction; or the mutual actions of two bodies upon each other are always equal, and directed to contrary parts."

<u>Components of the Third Law</u>

(1) The law applies to *two general bodies* (or particles)

a) Discussion must mention 2 bodies, and

b) These must be general bodies or particles
(Particular example bodies *alone* are not sufficient to meet this condition, although example bodies are allowed to be present)

(2) Action and reaction refer to *Forces* exerted by each body on the other, where these forces need not be of any particular type

a) Must be an explicit statement that *each* body (however body is discussed) exerts a "force" on the other; and

b) "Force" must be in general terms (particular example forces, such as kick, push, alone won't do although such examples are allowed to be present)

(3) Reaction (however stated) is *equal in magnitude*

(4) Reaction (however stated) is *opposite in direction*

(5) Line of action/reaction is in a *straight line between two bodies*

Table 7

Newton's Third Law Decomposed into Five Components and Two Sample Protocols

	Novice				Expert			
	K.D.	S.B.	J.W.	C.H.	O.G.	M.V.	S.D.	B.P.
Reaction opposite in direction	X	X	X	X	X	X	X	X
Reaction equal in magnitude	X		X	X	X	X	X	X
Action-Reaction involves two general bodies					X	X		X
Action-Reaction are general forces extended by each body on the other					X	X		X
Direction of Action-Reaction is a straight line					X			

Examples of Subjects' Summary Protocol

Nov. S.B. "And his third law states that for every action there's an opposite reaction to it."

Exp. O.G. "The third law. . . states that for every action there is an equal and opposite reaction, or in other words, if Body A exerts a force on body B, then Body B exerts a force on Body A in a direction which is along the line joining the two points. When you say bodies in this chapter, you mean they are really particles, point masses."

Table 8

Newton's First Law: "Every body persists in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by forces acting on it."

	Novice					Expert		
	J.W.	S.B.	K.D.	C.H.	S.D.	O.G.	M.V.	B.P
No Net Unbalanced Force	X		X	X	X	X	X	X
Rest		X	X					X
Uniform Motion				X	X	X	X	X
Straight Line						X	X	X

Examples of Subjects' Summary Protocol

Nov. J.W. "The first one is inertia, which is that a body tends to stay in a certain state unless a force acts upon it."

Nov. S.B. "First of all there's, the body wants to stay at rest, the body just, it's resistance toward any other motion."

Exp. B.P. "His first law is a statement that a body is moving in a uniform velocity in a given straight line or statics. It will keep moving or stay where it is unless some external forces are applied."

Exp. O.G. "The first law is called the law of inertia. And it states that a body persists in its motion along a straight line of a uniform rate unless a net unbalanced force acts upon the body."

Table 9

Expert Productions Converted from Protocols

M.S.

1. IF problem involves an inclined plane
   THEN a) expect something rolling or sliding up or down
   b) use F = MA
   c) use Newton's 3rd Law

*2. IF plane is smooth
   THEN use Conservation of Mechanical Energy

3. If plane is not smooth
   THEN use work done by friction

4. IF problem involves objects connected by string and one object being pulled by the other
   THEN consider string tension

5. IF string is not taut
   THEN consider objects as independent

M.G.

1. (IF problem involves inclined plane)[a]
   THEN a) use Newton's Law
   b) draw force diagram

*2. (IF problem involves inclined plane)[a]
   THEN can use Energy Conservation

3. IF there is something on plane
   THEN determine if there is friction

4. IF there is friction
   THEN put it in diagram

5. (IF drawing diagram)[a]
   THEN put in all forces - gravity, force up plane, friction, reaction force

6. (IF all forces in diagram)[a]
   THEN write Newton's Law's

7. IF equilibrium problem
   THEN a) $\Sigma F = 0$
   b) decide on coordinate axes

8. IF acceleration is involved
   THEN use F = MA

9. IF "that's done" (drawing diagram, putting in forces, choosing axes)[a]
   THEN sum Components of forces

---

[a] Statements in parentheses were not said explicitly by the subject but are indicated by the context.

Table 10

Novice Productions Converted from Protocols

H.P.

1. (IF problem involves inclined plane)[a]
THEN find angle of incline with horizontal

*2. If block resting on plane
THEN a) find mass of block
b) determine if plane is frictionless or not

3. IF plane has friction
THEN determine coefficients of static and kinetic friction

4. IF there are any forces on the block
THEN . . . . .

5. IF the block is at rest
THEN . . . . .

6. IF the block has an initial speed
THEN . . . . .

7. IF the plane is frictionless
THEN the problem is simplified

8. IF problem would involve Conservation of Energy and height of block, length of plane, height of plane are known
THEN could solve for potential and kinetic energies

P.D.

*1. (IF problem involves an inclined plane)[a]
THEN a) figure out what type of device is used
b) find out what masses are given
c) find outside forces besides force coming from pulley

2. IF pulley involved
THEN try to neglect it

3. IF trying to find coefficient of friction
THEN slowly increase angle until block on it starts moving

4. IF two frictionless inclined planes face each other and a ball is rolled from a height on one side
THEN ball will roll to same height on other side

5. IF something goes down frictionless surface
THEN can find acceleration of gravity on the incline using trigonometry

6. IF want to have collision
THEN can use incline to accelerate one object

[a] Statements in parentheses were not said explicitly by the subject but are indicated by the context.

Table 11

Key Features Cited by Experts and Novices

	Experts	
	V. V.	J. L.
Given initial conditions	9	3
Before and after situations	3	4
* Spring	0	5
No external force	4	1
Don't need details of motion	4	1
* Given final conditions	5	0
Asked something at an instant in time	4	1
* Asked some characteristics of final condition	4	0
* Interacting objects	0	4
* Speed — distance relation	0	4
* Inelastic collision	2	2
* No initial conditions	4	0
No final conditions	4	0
Energy easy to calculate at two points	1	2
No friction or dissipation	3	1
* Force too complicated	0	3
Momentum easy to calculate at two points	2	1
* Compare initial and final conditions	2	0
* Can compute work done by external force	2	0
Given distance	1	1
* Rotational component	0	2
* Energy yields direct relation	0	2
* No before and after	2	0
* Asked about force	2	0

	Novices	
	P. D.	J. W.
Friction	3	5
Gravity	3	3
Pulley	3	3
Inclined plane	3	2
Spring	2	3
Given masses	3	2
Coin on turntable	1	1
Given forces	1	1
* Force — velocity relation	0	2

* Asterisks indicate features mentioned by only one of the two subjects.

Table 12

Decomposition of a Problem Statement into Propositions

Problem No. 8

1. A block of mass M1
2. is put on top of a $\boxed{\text{block of mass M2}}$
*3. In order to cause the top block to slip on the bottom one,
*4. a $\boxed{\text{horizontal force}}$ F1 must be applied to the top block
5. Assume a $\boxed{\text{frictionless}}$ table.
*6. Find the $\boxed{\text{maximum horizontal}}$ force F2
7. which can be applied to the lower block
*8. so that both blocks will $\boxed{\text{move together}}$.

Table 13

Proportion of Response Types

	Novices	Experts
Abstract Principle	9%	30%
"straightforward application of Newton's Second Law"		
"collision problem, use Conservation of Momentum"		
"no friction, no dissipative forces, just apply Energy Conservation"		
Problem Characteristics	33%	35%
"frictionless, problem is simplified"		
"massless spring simplifies problem"		
"pulley introduces difficulty"		
Nonphysics Related Characteristics	40%	28%
"problem is difficult to visualize"		
"easy calculations but hard to understand"		
"many factors to consider, make problem difficult"		
Nonproblem related Characteristics	18%	7%
"never did problems like this"		
"numbers instead of symbols"*		
"must consider units"		
"diagram distracting"		

*all our problems used symbols for known quantities rather than actual numerical values

Table 14

Inferences Generated from Literal and Derived Cues

Literal Cue	Derived Cue	Inference
Frictionless		Conservation of Momentum
Frictionless	No dissipative forces	
	No dissipative forces	Conservation of Momentum
	No dissipative system	Conservation of Energy
Frictionless	No dissipative force	Conservation of Energy
Frictionless	No dissipative force	Conservation Laws
	Energy not consumed	Conservation of Momentum then calculate new Energy
Frictionless	Only force is restoring force	Newton's Second Law
Center of Mass at rest	No external forces	$\|M_1 V_1\| = \|M_2 V_2\|$
Center of Mass at rest		
Center of Mass at rest		Relative Momentum = 0
	Pulley must be taken into account	Newton's Second Law for translation and rotation
Mass and Radius of Pulley		Consider Rotational Kinetic Energy
	Pulley can't be neglected	Rotational Dynamics
Mass of Pulley		Rotational Energy
Massive Pulley		Rotational Dynamics
Compressing Spring		Think Energy
Motion		Energy Analysis
Slip and Force	Friction	
M_1 + M_2 Collide		Conservation of Energy and Momentum
M_2 Stops after distance L		Work-Energy
Speed		Newton's Second Law to Find Acceleration then Equation of Motion
Merry-Go-Round	Rotational Motion	Conservation of Angular Momentum

Navy

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Meryl S. Baker
NPRDC
Code P309
San Diego, CA 92152

1 Dr. Robert Breaux
Code N-711
NAVTRAEQUIPCEN
Orlando, FL 32813

1 Dr. Richard Elster
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Henry M. Halff
Department of Psychology,C-009
University of California at San Diego
La Jolla, CA 92093

1 LT Steven D. Harris, MSC, USN
Code 6021
Naval Air Development Center
Warminster, Pennsylvania 18974

1 Dr. Jim Hollan
Code 304
Navy Personnel R & D Center
San Diego, CA 92152

1 CDR Charles W. Hutchins
Naval Air Systems Command Hq
AIR-340F
Navy Department
Washington, DC 20361

Navy

1 CDR Robert S. Kennedy
Head, Human Performance Sciences
Naval Aerospace Medical Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 CAPT Richard L. Martin, USN
Prospective Commanding Officer
USS Carl Vinson (CVN-70)
Newport News Shipbuilding and Drydock Co
Newport News, VA 23607

1 Dr. James McBride
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr William Montague
Navy Personnel R&D Center
San Diego, CA 92152

1 Ted M. I. Yellen
Technical Information Office, Code 201
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 Library, Code P201L
Navy Personnel R&D Center
San Diego, CA 92152

1 Technical Director
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

Navy

1 Psychologist
ONR Branch Office
Bldg 114, Section D
666 Summer Street
Boston, MA 02210

1 Psychologist
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

5 Personnel & Training Research Programs
(Code 458)
Office of Naval Research
Arlington, VA 22217

1 Psychologist
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 Special Asst. for Education and
Training (OP-01E)
Rm. 2705 Arlington Annex
Washington, DC 20370

1 Office of the Chief of Naval Operations
Research Development & Studies Branch
(OP-115)
Washington, DC 20350

1 Dr. Donald F. Parker
Graduate School of Business Administrati
University of Michigan
Ann Arbor, MI 48109

1 LT Frank C. Petho, MSC, USN (Ph.D)
Code L51
Naval Aerospace Medical Research Laborat
Pensacola, FL 32508

Navy

1 Dr. Gary Poock
Operations Research Department
Code 55PK
Naval Postgraduate School
Monterey, CA 93940

1 Roger W. Remington, Ph.D
Code L52
NAMRL
Pensacola, FL 32508

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Roger Weissinger-Baylon
Department of Administrative Sciences
Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Wisher
Code 309
Navy Personnel R&D Center
San Diego, CA 92152

1 Mr John H. Wolfe
Code P310
U. S. Navy Personnel Research and
Development Center
San Diego, CA 92152

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Dexter Fletcher
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria,VA 22333

1 DR. FRANK J. HARRIS
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Michael Kaplan
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Training Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
Attn: PERI-OK
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasmor
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Frederick Steinheiser
U. S. Army Reserch Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Dr. Earl A. Alluisi
HQ, AFHRL (AFSC)
Brooks AFB, TX 78235

1 Dr. Genevieve Haddad
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Dr. Marty Rockway
Technical Director
AFHRL(OT)
Williams AFB, AZ 58224

2 3700 TCHTW/TTGH Stop 32
Sheppard AFB, TX 76311

Marines

1 H. William Greenup
Education Advisor (E031)
Education Center, MCDEC
Quantico, VA 22134

1 Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 N. Quincy St.
Arlington, VA 22217

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

Other DoD

12 Defense Technical Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC

1 Military Assistant for Training and
Personnel Technology
Office of the Under Secretary of Defense
for Research & Engineering
Room 3D129, The Pentagon
Washington, DC 20301

1 DARPA
1400 Wilson Blvd.
Arlington, VA 22209

Civil Govt

1 Dr. Susan Chipman
Learning and Development
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Joseph I. Lipson
SEDR W-638
National Science Foundation
Washington, DC 20550

1 William J. McLaurin
Rm. 301, Internal Revenue Service
2221 Jefferson Davis Highway
Arlington, VA 22202

1 Dr. Arthur Melmed
National Intitute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. Frank Withrow
U. S. Office of Education
400 Maryland Ave. SW
Washington, DC 20202

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 Dr. John R. Anderson
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Anderson, Thomas H., Ph.D.
Center for the Study of Reading
174 Children's Research Center
51 Gerty Drive
Champiagn, IL 61820

1 Dr. John Annett
Department of Psychology
University of Warwick
Coventry CV4 7AL
ENGLAND

1 DR. MICHAEL ATWOOD
SCIENCE APPLICATIONS INSTITUTE
40 DENVER TECH. CENTER WEST
7935 E. PRENTICE AVENUE
ENGLEWOOD, CO 80110

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

1 Dr. Patricia Baggett
Department of Psychology
University of Denver
University Park
Denver, CO 80208

1 Mr Avron Barr
Department of Computer Science
Stanford University
Stanford, CA 94305

Non Govt

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Lyle Bourne
Department of Psychology
University of Colorado
Boulder, CO 80309

1 Dr. John S. Brown
XEROX Palo Alto Research Center
3333 Coyote Road
Palo Alto, CA 94304

1 Dr. Bruce Buchanan
Department of Computer Science
Stanford University
Stanford, CA 94305

1 DR. C. VICTOR BUNDERSON
WICAT INC.
UNIVERSITY PLAZA, SUITE 10
1160 SO. STATE ST.
OREM, UT 84057

1 Dr. Pat Carpenter
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Charles Myers Library
Livingstone House
Livingstone Road
Stratford
London E15 2LJ
ENGLAND

1 Dr. William Chase
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

Non Govt

1 Dr. William Clancey
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, Ma 02138

1 Dr. Lynn A. Cooper
LRDC
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

1 Dr. Meredith P. Crawford
American Psychological Association
1200 17th Street, N.W.
Washington, DC 20036

1 Dr. Kenneth B. Cross
Anacapa Sciences, Inc.
P.O. Drawer Q
Santa Barbara, CA 93102

1 Dr. Hubert Dreyfus
Department of Philosophy
University of California
Berkely, CA 94720

1 LCOL J. C. Eggenberger
DIRECTORATE OF PERSONNEL APPLIED RESEARC
NATIONAL DEFENCE HQ
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A OK2

1 Dr. Ed Feigenbaum
Department of Computer Science
Stanford University
Stanford, CA 94305

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

Non Govt

1 Mr. Wallace Feurzeig
Bolt Beranek & Newman, Inc.
50 Moulton St.
Cambridge, MA 02138

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Alinda Friedman
Department of Psychology
University of Alberta
Edmonton, Alberta
CANADA T6G 2E9

1 Dr. R. Edward Geiselman
Department of Psychology
University of California
Los Angeles, CA 90024

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Marvin D. Glock
217 Stone Hall
Cornell University
Ithaca, NY 14853

1 Dr. Daniel Gopher
Industrial & Management Engineering
Technion-Israel Institute of Technology
Haifa
ISRAEL

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

Non Govt

1 Dr. Harold Hawkins
Department of Psychology
University of Oregon
Eugene OR 97403

1 Dr. James R. Hoffman
Department of Psychology
University of Delaware
Newark, DE 19711

1 Glenda Greenwald, Ed.
"Human Intelligence Newsletter"
P. O. Box 1163
Birmingham, MI 48012

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Steven W. Keele
Dept. of Psychology
University of Oregon
Eugene, OR 97403

1 Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Dr. David Kieras
Department of Psychology
University of Arizona
Tuscon, AZ 85721

1 Dr. Kenneth A. Klivington
Program Officer
Alfred P. Sloan Foundation
630 Fifth Avenue
New York, NY 10111

1 Dr. Stephen Kosslyn
Harvard University
Department of Psychology
33 Kirkland Street
Cambridge, MA 02138

Non Govt

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 Dr. Jill Larkin
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Alan Lesgold
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

1 Dr. Michael Levine
Department of Educational Psychology
210 Education Bldg.
University of Illinois
Champaign, IL 61801

1 Dr. Robert A. Levit
Director, Behavioral Sciences
The BDM Corporation
7915 Jones Branch Drive
McClean, VA 22101

1 Dr. Charles Lewis
Faculteit Sociale Wetenschappen
Rijksuniversiteit Groningen
Oude Boteringestraat
Groningen
NETHERLANDS

1 Dr. Erik McWilliams
Science Education Dev. and Research
National Science Foundation
Washington, DC 20550

1 Dr. Mark Miller
Computer Science Laboratory
Texas Instruments, Inc.
Mail Station 371, P.O. Box 225936
Dallas, TX 75265

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

Non Govt

1 Dr. Donald A Norman
Dept. of Psychology C-009
Univ. of California, San Diego
La Jolla, CA 92093

1 Dr. Jesse Orlansky
Institute for Defense Analyses
400 Army Navy Drive
Arlington, VA 22202

1 Dr. Seymour A. Papert
Massachusetts Institute of Technology
Artificial Intelligence Lab
545 Technology Square
Cambridge, MA 02139

1 Dr. James A. Paulson
Portland State University
P.O. Box 751
Portland, OR 97207

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 Dr. Martha Polson
Department of Psychology
University of Colorado
Boulder, CO 80302

1 DR. PETER POLSON
DEPT. OF PSYCHOLOGY
UNIVERSITY OF COLORADO
BOULDER, CO 80309

1 Dr. Steven E. Poltrock
Department of Psychology
University of Denver
Denver, CO 80208

1 MINRAT M. L. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 1328
D-53 BONN 1, GERMANY

Non Govt

1 Dr. Fred Reif
SESAME
c/o Physics Department
Universitv of California
Berkely, CA 94720

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 DR. WALTER SCHNEIDER
DEPT. OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61820

1 Dr. Alan Schoenfeld
Department of Mathematics
Hamilton College
Clinton, NY 13323

1 Committee on Cognitive Research
% Dr. Lonnie R. Sherrod
Social Science Research Council
605 Third Avenue
New York, NY 10016

1 Robert S. Siegler
Associate Professor
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

1 Dr. Edward E. Smith
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. Robert Smith
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

Non Govt

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. ALBERT STEVENS
BOLT BERANEK & NEWMAN, INC.
50 MOULTON STREET
CAMBRIDGE, MA 02138

1 David E. Stone, Ph.D.
Hazeltine Corporation
7680 Old Springhouse Road
McLean, VA 22102

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 Dr. John Thomas
IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

1 Dr. Douglas Towne
Univ. of So. California
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

1 Dr. J. Uhlaner
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Non Govt

1 Dr. Benton J. Underwood
Dept. of Psychology
Northwestern University
Evanston, IL 60201

1 Dr. Phyllis Weaver
Graduate School of Education
Harvard University
200 Larsen Hall, Appian Way
Cambridge, MA 02138

1 Dr. David J. Weiss
N660 Elliott Hall
University of Minnesota
75 E. River Road
Minneapolis, MN 55455

1 DR. GERSHON WELTMAN
PERCEPTRONICS INC.
6271 VARIEL AVE.
WOODLAND HILLS, CA 91367

1 Dr. Keith T. Wescourt
Information Sciences Dept.
The Rand Corporation
1700 Main St.

N